Virulence of the stem rust fungus and non-host resistance against stem rust

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Feng Li

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Dr. Melania Figueroa

Oct 2019

Acknowledgements

The decision to pursue a doctorate degree at the Department of Plant Pathology at University of Minnesota (UMN) has been a life-changing experience. I cannot imagine this adventurous and fruitful journey without the support, love, and guidance from many people along the way.

First and foremost, I would like to thank my mentor and advisor, Dr. Melania Figueroa, for her continuous patience, support, motivation and guidance during my training. Melania always encourages me to step out of my comfort zone, fueling my interest and passion for research and allowing me to grow both professionally and personally. Particularly, Melania has served as a role model for me as a female scientist while I navigate my own scientific journey. In addition, I really appreciate her time and commitment in helping me improve my writing and oral presentation skills, including the write-up of this dissertation. I could not ask for a better Ph.D. mentor.

Many thanks also to my committee members Drs. Peter Dodds, Brian Steffenson, Matthew Rouse and Candice Hirsch for their invaluable feedback and insightful advice. Their hard questions during committee meetings and exams motivated me to keep broadening and advancing my knowledge and helped me to reaching milestone and overall navigate graduate school. In particular, I thank Dr. Peter Dodds who has played an important role in my training and provided me with motivation and guidance for my research as well as his advice for my dissertation. I have learned a lot from him.

I also appreciate the help and support from members in the Figueroa's Laboratory for their help when I needed. Also, thanks to Drs. Marisa Miller and Kevin Silverstein for their patience and help to teach me bioinformatics skills. In addition, I have been very fortunate to have collaboration and support from wonderful scientists outside of UMN. I am very grateful to my collaborators Drs. Evans Lagudah, Peter Dodds, Ricky Milne, Michael Ayliffe for their tremendous support during a research internship at Commonwealth Scientific and Industrial Research Organization (CSIRO) in Australia. I am also very thankful for the help of Drs. Narayana Upadhyaya and Jana Sperschneider at CSIRO and Benjamin Schwessinger at Australia National University. Their contributions in the Ug99 genome assembly project were very valuable.

I am very grateful to be part of the positive, supportive and collaborative community at the Department of Plant Pathology, the University of Minnesota, that allows me to have opportunities to interact with extraordinary scientists around the world, enhance my professional development and make life-long friends. My Ph.D. training would not be possible without the generous funding support from Hatch project MIN-22-078, the MnDrive Global Food Ventures Fellowship and 2Blades Foundation.

In additional to these, I would not be able to complete my degree with a positive and happy attitude without the continuous spiritual and emotional support, love, company and help from my host family, Bob and Bonnie and all my friends. Last but not least, I cannot express more gratitude to my family. My parents and my brother always encourage me to follow my dreams and have always been there when I have concerns and doubts about myself or my decisions.

Abstract

Wheat stem rust is a destructive disease caused by the fungal pathogen *Puccinia graminis* f. sp. *tritici* (*Pgt*), which poses a significant threat to global wheat production. A highly virulent *Pgt* strain known as Ug99 emerged in Uganda in 1998 overcame an important disease resistance gene *Sr31.* Ug99 can infect 90% of wheat cultivars worldwide. Part of the Ph.D. research included in this dissertation focuses on creating genomic resources to study *Pgt* and understanding the underlying genetic differences that explains virulence evolution. The research also aims to discover novel genes in mediating resistance or susceptibility against *Pgt* in the grass species, *Brachypodium distachyon*. In Chapter 1, I present a comprehensive literature review of our current state of knowledge of cereal rust fungi, particularly *Pgt* in light of the recent evolution of new broadly virulent races. I also discuss current genomic resources and approaches to study virulence evolution of rust fungi. Finally, I present an overview of genetic disease resistance and its mechanisms, as the main strategy to mitigate the effect of *Pgt* in wheat production.

Chapter 2 describes the construction of the first *de novo* haplotype-phased genome assemblies of *Pgt*, including Ug99 and an Australian *Pgt* isolate 21-0. Importantly, a systematic comparison of both genomes shows that the Ug99 lineage emerged through a somatic hybridization event. This study also provides the first molecular demonstration that whole nuclear exchange at the vegetative stage contributes to the evolution of rust virulence in the field. One important goal within our scientific community is to develop novel and durable stem rust disease management strategies. Thus, another objective of my dissertation was to investigate non-host resistance against rust fungi to identify genes that

may enable the development of these approaches. Using a reverse genetics approach, Chapter 3 describes a pipeline that merges phenotypic and genotypic screenings to identify rust defense-associated genes in *Brachypodium distachyon*, which serves as a non-host to several cereal rust fungi. Here, I utilized a collection of T-DNA insertional lines of *B. distachyon* to characterize genes of interest in *Brachypodium*-rust interactions. Two candidate genes, a WRKY transcription factor and a sugar transporter, were identified to play a role in non-host resistance. Finally, I also attach appendices that refer to my contributions to four other independent publications, including a research manuscript that defines the non-host status of *B. distachyon* to *Puccinia coronate*, the causal agent of oat crown rust; a research publication of the cloning of the first effector gene in *Pgt*; a review and perspectives on *B. distachyon* as a donor of resistance against cereal rusts and finally one pathogen profile of *P. coronata*. Overall, research presented in this dissertation has important implications for crop protection and advances the field of biology and genomics in rust fungi. By creating suitable genomic resources to study *Pgt* and enabling effector discovery, this dissertation is deemed to contribute to rust resistance gene stewardship and minimize wheat losses due to stem rust epidemics.

Table of contents

List of tables

List of figures

Chapter 1. Introduction

Wheat production and rust fungi

The role of wheat production in food security and factors that affect crop productivity

Wheat (*Triticum* spp.) is the most widely grown crop around the world (FAO, 2017). In 2017, the wheat area harvested was 218 million hectares (ha), which was considerably higher than other important grains such as rice for which 167 million ha were harvested (FAO, 2017). Wheat provides nearly one fifth of daily calories consumed by humans, and therefore its role as a major staple food to large populations in the world is well-recognized (FAO, 2017). In addition to human consumption, wheat is also used for animal feed in many countries (USDA 2019). China, India, Russia, U.S.A. and the European Union (EU) are among the highest wheat-producing countries, and their economies depend heavily on the production of this grain (FAO, 2017) (Figure 1.1). The worldwide wheat production is nearly 730 to 760 million tonnes per year (2014 -2018), and the annual global trade value for wheat is approximately US $50 billion (CIMMYT, 2017; FAO, 2017). While the current wheat supply meets global market demands, the world's population is projected to grow to 9.7 billion people by 2050 (FAO) posing an enormous pressure to increase grain production. Given this, there is a pressing need to increase annual wheat production to 1 billion tonnes by 2050 to satisfy a wheat demand that is predicted to increase by 1.6% per year (Singh et al., 2016; Figueroa et al., 2018). Hence, breeding new wheat cultivars to

stabilize yield and improve its nutrition potential has been the focus of researchers and breeding programs at large.

Diseases caused by pathogenic fungi are among the most important biotic stresses facing wheat production and can inflict annual yield losses reaching up to 15 to 20% (Figueroa et al., 2018). Several of the most important foliar and floral diseases of wheat caused by fungi are the wheat rusts (*Puccinia* spp.), powdery mildew (*Blumeria graminis* f. sp. *tritici*), Fusarium head blight (FHB, *Fusarium graminearum)*, tan spot (*Pyrenophora tritici-repentis*), as well as Septoria tritici blotch (STB, *Zymoseptoria tritici*) (Figueroa et al., 2018).

Rust fungi that threaten wheat production

Rust fungi are major constraints to grain production. Efforts to lessen rust damage date back to the Roman times as exemplified by a festival called Robigalia, which was held to appease the god of rust Robigus in order to prevent grain losses due to rust infection (Roelfs, 1982). Rust diseases are caused by fungi of the group Pucciniales (also known as Uredinales), which are part of the Kingdom Fungi, Phylum Basidiomycota, Class Pucciniomycetes (Margulis and Chapman, 2009). The Pucciniales group accounts for a third of the Basidiomycota fungi and is estimated to have ~ 8,000 species from 125 genera, of which *Puccinia* represents the largest genus (Alexopoulos CJ. et al., 1996; Aime et al., 2017) including the three cereal rust fungi that affect wheat production. Based on host specialization, rust fungi are taxonomically classified below the species level by forma

speciales (f. sp.) (Anikster, 1984; Savile, 1984), which can be morphologically identical but differ in which plant species they are able to infect. Furthermore, physiological races, also referred to as pathotypes, can be defined within species or formae speciales based on examination of disease reactions (avirulence or virulence) in a set of host plant accessions that carry different resistance genes (Roelfs and Bushnell, 1984; Leonard and Szabo, 2005). Each race or pathotype would be capable of infecting a different combination of these differential host lines.

Puccinia graminis f. sp. *tritici* (*Pgt*) affects the production of bread wheat (*Triticum aestivum* L.), durum wheat (*Triticum turgidum* L. var. *durum*) and barley (*Hordeum vulgare* L.). *Pgt* is responsible for the disease wheat stem rust, also known as black rust or summer rust (Leonard and Szabo, 2005). After infection, massive numbers of rust-colored spores (urediniospores) erupt from the epidermis of stems, leaves or leaf sheaths, spikes, glumes and even grains. The infection by stem rust causes a reduction in the plant photosynthetic capacity and also a loss of nutrients in the host, resulting in grain shriveling. Also, *Pgt* infection reduces the strength of stems and sheaths which leads to plant lodging (Roelfs, 1992; Leonard and Szabo, 2005). Given these facts, stem rust can cause severe yield losses often ranging from 50% to 70% and could reach 100% under favorable environmental conditions (warm and wet) (Saari and Prescott, 1985).

Puccinia striiformis f. sp. *tritici* (*Pst*) causes stripe rust (yellow rust) of wheat. Aside from infecting wheat, this pathogen can also infect some barley and rye cultivars, but severe epidemics have not been reported in these crops (Chen, 2005; Chen et al., 2014). *Pst* forms stripes of yellow uredinia on wheat leaves, usually between veins. The fungus

can also infect spikes. Stripe rust is usually prevalent in temperate regions with a cool and moist environment, such as in China and Europe (Chen, 2005; Chen et al., 2014). Stripe rust currently causes the most economic losses of wheat among all three cereal rusts, resulting in 5% - 50% yield reduction under epidemic conditions (Singh et al., 2016).

Wheat leaf rust caused by *Puccinia triticina* (*Pt*) is the third rust disease of bread wheat and durum wheat. It is also known as brown rust (Bolton et al., 2008). Upon infection, brighter rust colored uredinia develop on the upper side of the leaf, and in cases of severe infection the fungus can also affect the sheaths of the wheat plants (Figueroa et al., 2018). *P. triticina* is the most widely adapted pathogen among all three wheat rusts as it is present in wheat growing regions with various climatic conditions (Kolmer, 2005; Singh et al., 2016).

It is estimated that wheat rust disease epidemics can lead to an economic loss of 4.3 to 5.0 billion US dollars in the wheat industry at the global level (Figueroa et al., 2018). Given the menace that *Puccinia* spp. represent to food security, these fungi have been ranked as the third most dangerous plant fungal pathogens (Dean et al., 2012). A common lifestyle feature of rust fungi is that they are obligately biotrophic, which means that they can only survive and reproduce by infecting living plants (Voegele and Mendgen, 2011). Their dispersal is aided by the airborne nature of the spores, which can travel long distances following the movement of air currents (Leonard and Szabo, 2005; Bolton et al., 2008; Hodson, 2011; Chen et al., 2014). In the recent decade, climate changes have created global concerns as the warming weather may reduce the effectiveness of resistance in wheat cultivars and favor rust growth and adaptation (Chakraborty et al., 2011). For example,

epidemiological models predict that wheat stem rust will develop in northern European countries such as Scotland where *Pgt* has never been reported as a significant issue due to cool weather (Davies K. et al., 2007). Also, *Pst* isolates and populations have adapted to warmer climates and decreased wheat yields in many other countries including the United States (Milus et al., 2009; Wellings, 2012b; Chen et al., 2014; Singh et al., 2016).

Given the economic importance of wheat rusts, it is fundamental that we advance our understanding of the biology of rust pathogens and disease resistance mechanisms to develop disease management strategies that lessen the impact of rust diseases. The research presented in this dissertation is focused on filling knowledge gaps in the areas of plant immunity and rust genome biology and providing support for wheat breeding and surveillance programs for stem rust.

Life cycle of the stem rust fungus *Puccinia graminis* f. sp. *tritici*

Puccinia graminis f. sp. *tritici* (*Pgt*) is a macrocyclic heteroecious rust (Leonard and Szabo, 2005). The complete life cycle of this fungus is complex and includes five spore types referred to as urediniospores, teliospores, basidiospores, pycniospores and aeciospores. As a heteroecious pathogen, *Pgt* requires two hosts to complete its life cycle. The asexual stage occurs on a gramineous host such as wheat and barley, and the sexual stage takes place on an alternate host such as *Mahonia* and *Berberis* species (Leonard and Szabo, 2005; Jin, 2011). It is important to note that the alternate hosts are distantly related to wheat and barley. *Mahonia* and *Berberis* are classified as dicot species and belong to

family Berberidaceae (Loconte and Estes, 1989), whereas wheat and barley are monocot plants in the Poaceae family (Soreng et al., 2015). Thus, the evolution of *Pgt* is rather extraordinary as it can infect species that are quite distinct.

Throughout the asexual phase of its life cycle, *Pgt* exists as a dikaryotic organism that contains two distinct haploid nuclei per cell with a full set of chromosomes within each nucleus (genome content n + n, where n represents one haploid genome equivalent). However, during the sexual stage of the life cycle, *Pgt* also exists as a monokaryotic haploid organism (n) after undergoing meiosis prior to sexual reproduction (Leonard and Szabo, 2005). The asexual phase of the *Pgt* life cycle begins when dikaryotic aeciospores (n + n) formed after sexual reproduction on barberry infect wheat plants. Subsequently, pustules (uredinia) erupt from the wheat leaf epidermis and release millions of dikaryotic urediniospores (n + n). Such high-level inoculum can then induce repeated secondary infections during the summer, which can occur every fourteen days depending on weather conditions and susceptibility of neighboring wheat plants. As infected plants mature and senesce, the fungus produces black teliospores, which are overwintering structures. Although teliospores are initially dikaryotic (n + n), the two nuclei in each teliospore eventually fuse (karyogamy) to produce one single diploid nucleus (2n) with two sets of chromosomes. In the spring, the teliospores (2n) germinate into a basidium that produces four haploid basidiospores (n) through meiosis. Basidiospores (n) are transported by wind to neighboring barberry bushes to complete the sexual life cycle. This haploid stage infection results in the formation of pycnia on the upper surface of a barberry leaf, which produce both free pycniospores (n) as well as receptive hyphae. Pycniospores transferred

between pycnia are able to fuse with the receptive hyphae (plasmogamy) of a compatible mating type, resulting in the development of dikaryotic aeciospores (n + n) on the lower surface of barberry leaf. Later, aeciospores are transported to wheat plants to cause primary infection. If the alternate host is absent, *Pgt* can propagate solely by clonal reproduction of urediniospores in wheat indefinitely without completing a sexual reproductive phase (Agrios, 1997; Leonard and Szabo, 2005).

The infection process of *Puccinia graminis* f. sp. *tritici* on wheat

Wheat infection by *Pgt* begins when urediniospores land on the leaf surface and germinate to form germ tubes in the presence of free water (Leonard and Szabo, 2005). Germ tubes grow on the leaf surface guided by topographical signals (thigmotropism), and differentiate a structure called an appressorium when they encounter a stoma. The appressorium forms a penetration peg that grows between the stomatal guard cells. Subsequently, a substomatal vesicle differentiates within the leaf cavity. Primary infection hyphae branch out from this vesicle and grow intercellularly in the plant apoplast. Once in contact with mesophyll cells, hyphae differentiate haustorial mother cells that form feeding structures known as haustoria to extract nutrients from plant cells. The haustorium penetrates the plant cell wall and invaginates the plant plasma membrane. Even though a haustorium develops within a living host cell, it remains separate from the plant cytoplasm through a complex haustorium-host interface (Staples, 2001; Garnica et al., 2014). The haustorial membrane is surrounded by a fungal cell wall, an extrahaustorial matrix, and finally the plant-derived exhaustorial membrane. This interface is considered to serve as a

"channel" or "bridge" via which nutrient transport (sugar and amino acids) or effector delivery occurs.

Environmental factors such as the presence of light, the level of CO_2 and the degrees of water stress can influence the stem rust infection process (Yirgou and Caldwell, 1963; Burrage, 1970). Previously, it was believed that penetration is light dependent because light exposure facilitates stomatal opening (Hart, 1929). However, a recent microscopy-based study using highly sensitive and less disruptive detection methods for visualizing *in planta* fungal growth discovered that penetration in the case of stem rust is light independent (Solanki et al., 2019). In contrast, intercellular pathogen growth, branching and haustoria formation are light dependent (Solanki et al., 2019).

Molecular basis of rust virulence

Pgt virulence is governed by similar molecular events that dictate virulence in other plant pathosystems. A successful infection depends on whether the plant is a good host and how well the microbe is capable of colonizing the plant (van Schie and Takken, 2014). The pathogen must gain entry to the host environment by overcoming the preformed obstacles of plants including both physical and chemical barriers (Freeman and Beattie, 2008). Examples of these defenses are deposition of cell wall appositions (Hückelhoven, 2007) and production of antimicrobial compounds (González-Lamothe et al., 2009). If pathogens successfully achieve entry into the plant, their growth and multiplication within the host relies on interference with the plant immunity system. Inducible defense mechanisms are

triggered by pathogen recognition through various plant receptors at different locations in a plant cell (Dodds and Rathjen, 2010). In general, plant immunity can be classified into two branches based on the nature of pathogen recognitions. The first branch of recognition involves pattern-recognition receptors (PRRs) such as receptor-like kinases. These receptors can recognize highly conserved molecules from a class of pathogens referred to as pathogen-associated molecular patterns (PAMPs), i.e. chitin in fungal cell walls. PRRs are localized in the plasma membrane of the plant cell, and therefore PAMPs are extracellularly recognized. Subsequently, the recognition of these molecules activates a cascade of defense signaling networks intracellularly, which results in the induction of PAMP-triggered immunity (PTI) (Dodds and Rathjen, 2010). However, adapted pathogens have evolved mechanisms to disarm PTI via the secretion of effector proteins/molecules into the plant cytoplasm. Effectors act as manipulators to alter host structures or defense signaling to support pathogen virulence and fitness (Jones and Dangl, 2006; Dodds and Rathjen, 2010). In many pathosystems, some effectors (also referred to as Avr proteins) are recognized by intracellular nucleotide-binding leucine-rich repeat receptors (NLRs) (Dodds and Rathjen, 2010). This recognition activates effector-triggered immunity (ETI) which constitutes the second branch of plant immunity (Ayliffe et al., 2010; Dodds and Rathjen, 2010). To date, we understand that a physiological race of a particular rust strain is genetically determined by the repertoire of individual effectors encoded by its genome and the host NLR (*R*) genes that recognize those effectors (Dodds and Rathjen, 2010; Figueroa et al., 2016). Effectors that escape recognition by NLRs in the plant define compatibility and host range of the pathogen. These molecular interactions are the

foundation for a co-evolutionary arms race between plants and pathogens (Jones and Dangl, 2006).

Rust fungi are predicted to encode hundreds of effectors (Upadhyaya et al., 2015; Figueroa et al., 2016). Contrary to bacterial systems (Toruño et al., 2016), we know very little about the identity and functions of rust effectors. As learned from the flax rust model system (*Melampsora lini*), effectors are secreted from rust haustoria and translocated into plant cells by a mechanism that has not been elucidated yet (Garnica et al., 2014). In all cases it appears that the defense responses associated with ETI are triggered by a direct interaction of NLR proteins in flax and corresponding Avr proteins from *M. lini* (i.e. L6, L7 and Avr567) (Dodds et al., 2006). As illustrated by the flax rust fungus, effector genes are highly polymorphic and often show signatures of diversifying selection (Figueroa et al., 2016). It is thought that the direct interaction of effectors with NLRs imposes strong divergent selection driving changes in DNA sequences that allow the effectors to escape recognition (Dodds et al., 2006; Barrett et al., 2009). However, not all pathogen Avr proteins are detected by direction interactions. The "guard hypothesis" and "decoy hypothesis" are examples of models of indirect recognition between effectors and NLRs (DeYoung and Innes, 2006; Cesari et al., 2014). Although no evidence of indirect recognition of effectors in rust fungi has yet identified, this process is known to occur in bacterial disease systems. Perhaps future research in rust biology will find evidence of indirect recognition of effectors.

While significant progress has been made in the identification and functional characterization of effectors in flax rust, the molecular mechanisms of virulence in cereal

rust fungi remain elusive. So far only one segregating population has been generated to identify markers linked to avirulence genes in *Pgt* (Zambino et al., 2000) and could serve as a resource to enable positional cloning of effectors of *Pgt*. Considering the difficulty in both setting up sexual crosses in *Pgt* and building a substantial mapping family, researchers have focused on identifying spontaneous mutants of *Pgt* or developed methodologies to induce mutations via chemical mutagens or radiation (Luig, 1978; Gates and Loegering, 1991) as a way to create genetic resources for isolation of *Avr* genes. *Avr* gene identification and characterization of allelic variation at these loci is critical to inform breeding decisions and develop sustainable strategies for *R* gene deployment. Work presented in this dissertation (Appendix 3) was critical for the identification of the first *Avr* gene in *Pgt*, *AvrSr50* (Chen et al., 2017).

Stem rust epidemics and pathogen evolution

Stem rust as an airborne pathogen has caused severe epidemics in wheat fields

Historically, stem rust has been a threat to many countries growing wheat. During infection, a common sign of *Pgt* is the formation of clusters of spores, known as pustules (Leonard and Szabo, 2005). Each pustule can produce ~100,000 urediniospores daily, and one in ten of these can develop successful infection under favorable conditions in the field (Katsuya and Green, 1967). The severity of epidemics may vary across fields depending on the type of the cultivars, infection time and environmental conditions (Eversmeyer and Kramer, 2000). Dispersion of the stem rust pathogen to neighboring wheat fields occurs

via urediniospores spread by wind, rain, animals or humans (Wellings, C.R. et al., 1987; Viljanen and Cromey, 2002; Singh et al., 2008). Even healthy wheat crops approaching harvest can be rapidly destroyed by stem rust infection when enough inoculum is transferred from a distant heavily infected crop (Leonard and Szabo, 2005; Singh et al., 2015; Visser et al., 2019). Extensive epidemiological studies showed the role of wind-blown urediniospores in the initiation and spread of stem rust epidemics (Kolmer and Peterson, 2001). In addition, field surveys, global meteorological data and computational simulations have also demonstrated that stem rust is a true global threat given the long distance transport of rust urediniospores (Meyer et al., 2017a, 2017b; Visser et al., 2019). In fact, intercontinental dispersal of *Pgt* isolates by high-altitude winds has been documented several times (Watson and Sousa, 1983; Luig, 1985). For example, surveillance of stem rust suggests that four exotic introductions occurred in Australia from South and East Africa since 1921 (Watson and Sousa, 1983; Park, 2007; Visser et al., 2019), including the race 21 which migrated from South Africa and arrived in Australia in 1954 (Park, 2007; Visser et al., 2019). These claims have been supported by recent genetic analysis and atmospheric dispersion modeling assays (Visser et al., 2019).

Two of the most destructive epidemics of stem rust in the United States occurred in 1935 and 1953-1954, with these outbreaks resulting in over 50% and 75% yield losses, respectively (Stakman and Harrar, 1958; Roelfs, 1978). Since then, the most significant worldwide epidemic of stem rust was caused by Ug99 in Africa. Ug99 was first detected in Uganda in 1998 (Pretorius et al., 2000) and later designated as race TTKSK based on the North American nomenclature system (Wanyera et al., 2006). This nomenclature is a

"letter code" system used to define different physiological races of *Pgt*. The system currently employs five sets of four single-gene differential wheat lines, with each set designated as a letter based on the sequence of infection scores of plants (High infection types: H or Low infection types: L) (Roelfs and Martens, 1988; Jin et al., 2008). Ug99 is a broadly virulent isolate of the stem rust fungus which overcomes most of the disease resistance genes deployed in wheat cultivars. Since its first detection, this strain of *Pgt* has spread across eastern and southern Africa, and 12 additional variants has been produced through continual asexual (clonal) reproduction, together forming the Ug99 lineage (Pretorius et al., 2012; Fetch et al., 2016; Patpour et al., 2016; Terefe et al., 2018). The Ug99 lineage has caused large-scale wheat yield losses in East Africa, South Africa and the Middle East as they have spread through these regions via wind trajectories (Singh et al., 2015).

Several other outbreaks of stem rust due to non-Ug99 races have occurred in Africa and Europe in recent years. For instance, the 'Digalu'-infecting race (TKTTF) emerged in Ethiopia in 2014 and overcomes the *SrTmp* resistance gene present in the 'Digalu' cultivar, which provided resistance against Ug99 (Olivera et al., 2015). The 'Digalu' race has spread to the Middle East, East Africa, Germany and the UK (Olivera et al., 2015, 2017; Singh et al., 2015; Lewis et al., 2018). Another example is the highly virulent *Pgt* strain which belongs to race TTTTF which emerged in Sicily, Italy in 2016. This race led to the largest stem rust outbreak in Europe in the past several decades (Bhattacharya, 2017).

In summary, the presence of *Pgt* races with different virulence profiles, including the Ug99 lineage, poses significant challenges for breeding schemes guarding against

large-scale rust epidemics. Therefore, international collaboration on disease surveillance of stem rust and deployment of effective resistance genes are critical components to battle this "shifty enemy" which is capable of traveling long distances (Kolmer and Peterson, 2001; Ellis et al., 2014).

The clonal lineage of *Pgt* Ug99

Dr. Norman Borlaug, architect of the Green Revolution, bred and deployed semi-dwarf rust-resistant wheat cultivars for the prevention of stem rust epidemics worldwide (Borlaug, 1968). This endeavor was successful for nearly half a century until the emergence of the *Pgt* race Ug99. Ug99 has captured intense interest in rust research, since this novel race and its lineage overcomes the wheat resistance gene *Sr31*, which is present in 90% of wheat cultivars grown worldwide (Pretorius et al., 2000; Singh et al., 2011). Ug99 has evolved rapidly since its discovery, and thirteen variants with different virulence pathotypes (TTKSK, TTKSF, TTKST, TTTSK, TTKSP, PTKSK, PTKST, TTKSF+, TTKTT, TTKTK, TTHSK, PTKTK and TTHST) have been detected in 13 countries in Africa, the Middle East and Asia (Pretorius et al., 2012; Fetch et al., 2016; Patpour et al., 2016; Terefe et al., 2018). The Ug99 lineage is likely to continue evolving virulence on deployed resistance genes and increasing its geographical expansion at regular intervals (Singh et al., 2015).

To date we know that the Ug99 lineage has overcome 35 stem rust resistance genes, including commercially important genes such as *Sr31* (Singh et al., 2015). Thus, there is a

clear need to identify new genetic resistance or understand mechanisms of plant-rust interactions that can strengthen current approaches to reduce or eliminate the negative effects of *Pgt* such as the Ug99 lineage. The work presented in this dissertation addresses some of these needs.

Genetic diversity and virulence evolution mechanisms of *Puccinia graminis* f. sp. *tritici*

Sexual reproduction of the stem rust fungus on common barberry allows this pathogen to develop races with new virulence combinations through meiotic recombination (Roelfs, 1982; Jin, 2011). This provides an opportunity for evolution of *Pgt* to increase the number of pathogenic races, while the production of aeciospores on barberry also provides an additional source of virulent inoculum in spring, both facilitating disease onset (Roelfs, 1982). In view of these observations, the federal government of the United States implemented a program on barberry eradication for decades which began in 1918 (Roelfs, 1982), which significantly decreased the frequency of stem rust epidemics. Likewise, barberry is not present in mainland Australia probably since the species was never introduced into the continent. *Pgt* has still evolved new virulences in this continent, established various clonal lineages and defeated several important resistance genes for the wheat production of Australia (Park, 2007). However, an unresolved critical question is how this pathogen gains genetic diversity especially in the absence of the alternate host and without the opportunity for sexual reassortment of genes.

Gene flow and somatic mutations

Gene flow and somatic mutations can provide clear paths for increasing genetic variation of rust fungi in the absence of sexual reproduction (Park, 2007; Singh et al., 2015). In the case of *Pgt* populations in Australia, several foreign introductions of rust strains provide classical examples of introduction of genetic diversity by gene flow (Park, 2007). Subsequently, the adaptation of *Pgt* isolates is proposed to have occurred by sequential random mutation of avirulence genes leading to the defeat of single resistance genes. Similarly, the Ug99 race group in Africa appears to be evolving through clonal reproduction and stepwise mutation (Singh et al., 2015). Race TTKSF (avirulent on *Sr31*) or PTKSF most likely represent the progenitor of Ug99 (TTKSK), which gained virulence on *Sr31* (Singh et al., 2015). Another race TTKSF+ may have emerged by a single-step mutation from TTKSF to become virulent on *Sr9h* (Singh et al., 2015). However, the origin of the Ug99 lineage and other *Pgt* clonal lineages is a long-standing question that has not yet been solved. One possibility is that these diverse groups each have their origin in a sexual reproduction event on an alternate host, which gave rise to a unique genotype that subsequently expanded clonally. Although *Berris holstii*, a species of the barberry family, may serve as an alternate host in Ethiopia, its geographic distribution appears to be limited, and its contribution to the evolution of *Pgt* has not been demonstrated. Chapter 2 of this dissertation presents evidence that the Ug99 lineage did not arise via sexual reproduction, but rather via somatic hybridization, which is further discussed below.

Current status of understanding the contribution of somatic hybridization to the evolution of Puccinia graminis* f. sp. *tritici

In addition to sexual recombination, other more cryptic processes, such as somatic recombination and somatic hybridization during the asexual stage of the life cycle of rust species, have been postulated to introduce genetic variability to populations (Park and Wellings, 2012). The occurrence of somatic hybridization within or between rust species has been discussed. However, most of these claims were based solely on phenotypic data or very limited genotypic data as these studies were conducted when molecular biology techniques were not available or just had started to be utilized. Below I review these past claims.

Early studies suggested that native and cultivated grasses could be implicated in the somatic hybridization between *Pgt* and *Puccinia graminis* f. sp. *secalis* (*Pgs*, the causal agent of rye stem rust) (Luig and Watson, 1972). In this case, the putative somatic hybrid, referred to as the scabrum rust, was detected in the wheat grass *Agropyron scabrum,* which served as a mutual host for *Pgt* and *Pgs* (Burdon et al., 1981). In addition, a few other studies have also suggested that somatic hybridization can occur within the same rust species in all three cereal rusts (Knott, 1989). In these cases, researchers noted that when inoculating a mixture of races within one rust species onto wheat seedlings, rust races with new virulence phenotypes different from the two co-inoculated (parental) strains were recovered (Little and Manners, 1969; Sharma and Prasada, 1969). For example, Sharma and Prasada inoculated seedlings of wheat cultivars with six pairs of two distinct *Pgt* pathotypes, and their analysis concluded that two pairs of these combinations produced

somatic recombinants (hybrids) given that the number of new virulent races was very high (Sharma and Prasada, 1969). The frequency and success of the production of the new races from some parental combinations over others was thought to be determined by somatic compatibility factors varied between parental isolates (Sharma and Prasada, 1969).

Other laboratory studies under controlled conditions also postulated that somatic hybridization could occur within and across formae speciales of cereal rust fungi. Watson and Luig (1958) co-inoculated wheat plants with two single spore cultures from two Australian field strains, orange 21-1 and red 222-2 pathotype (two different races) and obtained 6 isolates with virulence phenotypes different from the parental strains. One of these had the pathotype 21-2 which was the same as a field isolate that was collected in Australia (Watson and Luig, 1958). Thus, the authors claimed that somatic hybridization potentially occurred in the field. However, the interpretation of these results lack support, and one would think that the presence of a similar isolate in the field is only indirect evidence that somatic hybridization could actually take place in the field. It is possible that the new pathotypes had emerged from stepwise mutations of single virulence genes. Moreover, the race 21-2 was first identified on barberry in Tasmania and was hypothesized to have appeared from selfing of the prevalent race 21-1 on the island (Watson and Luig, 1958). Thus, it is possible that movement of spores could also explain the presence of 21-2 in Australia's mainland.

Only a few studies have supported claims of somatic hybridization with genotypic data and molecular methods. Burdon and colleagues (Burdon et al., 1982) utilized an isozyme marker analysis of eight enzymes to postulate that the *Pgt* isolate 34-2,11 was

most likely a hybrid of race 126 and race 21. In this study, all isozyme banding patterns of 34-2,11 resembled race 21 except for the glutamate oxalo-transaminase (GOT) which shared the unique pattern of race 126. Thus, the claim that the *Pgt* isolate 34-2,11 is a hybrid was only based on a single marker. Likewise, isozyme analyses suggested that the scrabrum rust isolates (putative somatic hybrids of *Pgs* and *Pgt*) had the identical leucine amino-peptidase (LAP) and NADH diaphorase (NADHD) pattern to *Pgs* and *Pgt*, respectively (Burdon et al., 1981). Again, although consistent with a hybrid origin, this interpretation is based on a single marker and does not rule out other potential mechanisms including mutation. In a more recent study, Lei and co-authors (Lei et al., 2017) identified new recombinant variants obtained after co-inoculation of isolate pairs of *P. striiformis*. Here, Lei et al. concluded that these variants were generated by somatic recombination on the basis of 51 simple-sequence repeat (SSR) markers and 90 single-nucleotide polymorphism (SNP) markers. It is important to note that the vast majority of claimed recombinants were almost identical to one of the parents, with only one or two markers apparently derived from the other parent. Such results were not expected from a somatic hybridization and the authors did not exclude the trivial explanation that these recombinants represented contaminant isolates or mutants. Overall the methods employed in these studies lack the necessary resolution to unequivocally support conclusions of somatic hybridization.

In the past, some studies of somatic hybridization were discredited as researchers later found out that the original isolates were not pure (Kolmer and Peterson, 2001). Thus, critical steps when interpreting results of somatic hybridization from virulence

phenotyping experiments include ruling out potential contamination from other rust isolates, any possibility of the presence of sexuality and meiotic recombination cycle of rust pathogens, as well as single step mutations. All these events could explain the recovery of isolates with different virulence profiles.

Some studies have suggested the somatic hybrids arose by parasexuality and somatic recombination, which involves formation of a diploid nucleus via nuclear fusion in vegetative cells followed by occasional mitotic crossing-over and vegetative haploidization (Knott, 1989). However, the occurrence and frequency of these events in rust fungi in nature remain unknown. This is due to 1) the lack of isolate- or race-specific molecular markers, which has hindered the critical examination of somatic hybridization/recombination particularly in nature; and 2) under field conditions, distinguishing between sexual recombination and somatic recombination seems impossible. Ruling out the possibility of a sexual origin could be challenging since this requires the availability of the genetic data of both haploid nuclei in the putative somatic hybrid and potential parents for comparisons to determine the absence of sexual recombination and the attribute of the hybridization event.

While the evidence presented by virulence traits, isozyme analysis and molecular markers is limited or inconclusive, it is reasonable to think that rust fungi may utilize somatic hybridization as an alternative evolutionary mechanism for the purpose of enhancing genetic diversity as parasexuality has been well documented for filamentous Ascomycete fungi (Burdon and Silk, 1997; Park and Wellings, 2012; Fleißner and Serrano, 2016; Stukenbrock, 2016). A simple somatic hybridization event would involve the fusion

of germ tubes or dikaryotic hyphae, followed by exchange and/or reassociation of whole nuclei (heterokaryosis between haploid nucleus with compatible mating types) (Burdon and Silk, 1997). The observation of these phenomena was described in *P. triticina* in a microscopy-based study (Wang and McCallum, 2009). According to this study, germ tube anastomosis resulted in the formation of a structure called a germ tube fusion body (GFB). These GFBs may provide the opportunity for nuclei to be exchanged between different rust hyphae or possibly recombine.

To date, we still lack robust molecular and genetic data to support the occurrence or underlying mechanisms of somatic hybridization and its consequences in the evolution of rust fungi. At this point no strong evidence supports that either somatic hybridization or recombination have important epidemiological effects. It is also unknown if rust hybrids have a high survival rate or suffer from any fitness penalties, which certainly would affect their frequency. In summary, there are key gaps in our understanding of evolutionary mechanisms that enrich genetic diversity of rust fungal populations, particularly the role of somatic hybridization in the occurrence of new *Pgt* isolates. As the community works together to identify new sources of genetic resistance to control diseases like stem rust, another clear research priority is to understand the genetic plasticity of rust fungi and factors contributing to virulence evolution as this is pivotal to minimize the evolutionary capacity of these pathogens. Chapter 2 of this dissertation is dedicated to delivering genomic resources for *Pgt* and advancing our understanding of the mechanisms that drive the evolution of rust fungi without sexual reproduction.

Genomics resources for *Puccinia graminis* f. sp. *tritici*

Overview of the existing genome references of *Pgt*

In the absence of genetic transformation techniques that could be adopted to study cereal rust fungi, the field of genomics has brought valuable opportunities to expand our knowledge of the biology of *Pgt*. Sequencing technology advances in the mid-2000s allowed the generation of reference genome sequences for various rust fungi, including the three cereal rusts and poplar leaf rust (*Melampsora larici-populina*) (Aime et al., 2017). *Pgt* was the first representative to be sequenced in the cereal rust fungi group (Duplessis et al., 2011). This *Pgt* reference genome was generated from genomic DNA extracted from urediniospores of the strain CRL 75-36-700-3 (race SCCL, 7a) followed by Sanger whole-genome shotgun sequencing of two plasmid libraries and one fosmid library, which allowed the assembly of a genome reference for the species. The genome assembly consisted of 4,557 contigs with a total size of ~81.5 Mbp and a contig N_{50} length of 39.5 kbp. A fingerprint map constructed in the fosmid library was used to scaffold these contigs, yielding a haploid genome assembly size of 88.6 Mbp, in 392 scaffolds, with about 7 Mbp in gaps between contigs (Duplessis et al., 2011). The genome of *Pgt* is highly repetitive and ~45% of its content corresponds to transposable elements (TEs). *Ab initio* and homology-based genome prediction and annotation yielded 17,773 protein coding genes. Subsequent genomic studies in *Pgt* have been based on short-read sequencing such as Illumina paired-end sequencing and focused on detecting intraspecific variation of *Pgt* isolates (Duplessis et al., 2014).

A pan genome of Australian *Pgt* isolates was generated in 2015 (Upadhyaya et al., 2015). The Australian *Pgt* isolate 21-0 was sequenced using Roche 454 technology and assembled using the *Pgt* 7a assembly as a reference. This effort yielded a genome assembly of 79.2 Mbp and the unmapped reads were utilized for *de novo* assembly that added an additional 13.3 Mbp. The additional ~13 Mbp novel sequences in the assembly of 21-0 are not necessarily isolate-specific as they could represent the missing sequences in the *Pgt* 7a assembly and the majority of these sequences are also present in other Australian isolates. As a result of these genome reference improvements, the total genome size of isolate 21-0 was estimated as ~92.5 Mbp. Next, the *Pgt* 21-0 reference was used to build a pan genome (PGTAus-pan) of Australian *Pgt* isolates, together with isolates representing four independent exotic introductions of *Pgt* into the continent. Adding to this, the researchers also included a transcriptome analysis from germinated spores and haustoria which detected ~3,500 additional new transcripts including potential alternative transcripts compared to the initial *Pgt* 7a genome reference. To determine the genomic diversity of *Pgt* between isolates, Illumina reads of each strain were aligned to the PGTAus-pan genome to detect variants. This analysis indicated a high level of heterozygosity reflected by 13 - 15 variants/kbp, with single-nucleotide polymorphisms (SNPs) and multiple-nucleotide polymorphisms (MNPs) representing 86% of the total variation (Upadhyaya et al., 2015).

Given the fact that *Pgt* genome contains two distinct nuclei (dikaryons), all previously available *Pgt* assemblies are actually mosaic haploid representatives of a dikaryon and do not capture all genetic information from both nuclei. Missing information

about haplotype variation within individual rust isolates presents obstacles for studying effector diversity in rust populations and distinguishing allelic variation of *Avr* genes at a high resolution (Figueroa et al., 2016).

Progress in phasing haplotypes of cereal rust genomes

Because rust fungi contain two separate haploid nuclei, an important goal is to try to assemble these genomes separately in order to define the level of diversity between them. The first attempts at such haplotype phasing of rust genomes were published in 2018 (Miller et al., 2018; Schwessinger et al., 2018). These studies performed long-read sequencing on a PacBio RSII sequencer to obtain data for *Puccinia coronata* f. sp. *avenae* (*Pca,* the causal agent of oat crown rust) and *P. striiformis* f. sp. *tritici* (*Pst*). Haplotype-resolved genome assemblies were generated employing the diploid-aware assembler FALCON and FALCON-Unzip. Miller et al (2018) sequenced two *Pca* isolates, 12SD80 and 12NC29, and generated *de novo* genome assemblies in which the total lengths were ~100 Mbp from primary contigs, approximately constituting a haploid genome size of the species. In addition, 52% of the genome of *Pca* was separated into alternate haplotypes, suggesting that haplotypes within a single individual isolate could be highly divergent. Consistent with this, more than 2% of the total genome size was impacted by large structure variation, and over 200,000 heterozygous SNPs were found within each isolate.

Schwessinger and co-authors generated a 83 Mbp genome assembly of *Pst* represented by 156 primary contigs with corresponding haplotype blocks (475 contigs)

representing another 73 Mbp (Schwessinger et al., 2018). The two haplotypes of *Pst* were even more divergent than those of *Pca*, with large structural variants accounting for ~6% of the genome. From an evolutionary perspective, the higher interhaplotype diversity of the Australian *Pst* isolate could be explained by the fact that *Pst* exclusively undergoes asexual reproduction and therefore lacks chromosome recombination of sexual stage, which can purge accumulated mutations. On the contrary, *Pca* isolates in the United States undergo frequent sexual reproduction cycles which may explain the higher similarity between their two haploid genomes (Miller et al., 2018; Schwessinger et al., 2018).

While the assemblies of *Pca* and *Pst* are significant advances in the field of genomics of cereal rust fungi, they still do not represent fully haplotype-resolved genome references, as the alternate haplotype sequences have not been assigned to whole nuclear haplotypes. The availability of long-read PacBio sequencing has the potential to overcome challenges of fully separating haplotypes. Firstly, long-read sequencing enables a more contiguous genome assembly for this large, repeat-rich fungal genome. Secondly, long-read and/or diploid-aware assembly is promising to separate genetic information from each karyon. Thirdly, long-read sequencing has the added benefits of detecting more genomic variants, especially large structural variation that could play a critical role in virulence evolution of effectors. In view of all these concepts, work presented in Chapter 2 of this dissertation has fully resolved the nuclear haplotypes of two *Pgt* isolates using PacBio-based sequencing technologies and new genome assembly algorithms along with genome comparison and DNA-proximity analysis.

Stem rust disease management and genetic resistance

Rust fungi are often described as "shifty enemies" due to the fact that these pathogens are constantly evolving new virulences (Stakman, 1947). This characteristic has posed a challenge to disease management particularly when farmers need cost-efficient and sustainable approaches. One common method to prevent yield losses caused by *Pgt* is the timely application of foliar fungicides; however, multiple applications of fungicides over the growing seasons can increase costs, enhance the risk of developing fungicide resistance, and raise environmental and human health concerns (Loughman et al., 2005; Wanyera et al., 2009; Pardey et al., 2013; Oliver, 2014).

To address these issues, scientists have focused on a more sustainable and environmentally friendly strategy which relies on employing genetic resistance against rust pathogens including *Pgt*. Breeding and deployment of new wheat cultivars that carry disease resistance genes against *Pgt* is effective for both reducing disease severity and mitigating stem rust epidemics (Ellis et al., 2014). Resistance genes of wheat are classified based on their race-specificity, durability and the plant stage when the resistance is effective, and these will be discussed in the following sections.

Race-specific resistance and interaction mechanisms

The majority of resistance genes that are deployed by breeding programs to protect against rust damage confer race-specific resistance, also known as major resistance or seedling resistance. As the name implies, these genes provide resistance against specific

races/strains of pathogen. The underlying principles of race-specific resistance are explained by the "gene-for-gene" concept proposed by H. Flor (Flor, 1971; Ellis et al., 2014), which stated that a resistance (*R*) gene in plants only confers resistance against pathogen strains carrying the corresponding avirulence gene (*Avr*) encoding effectors. Therefore, the effectiveness of race-specific resistance is mediated by R proteins present in a host cultivar and the array of effectors expressed by a pathogen race. To date, it is known that such recognition events activate host defense responses and usually result in a hypersensitive response (HR) (Dodds and Rathjen, 2010).

R proteins of the NLR class contain a nucleotide binding site (NBS) and leucine rich repeat (LRR) domain (Dodds and Rathjen, 2010). The C-terminal LRR domain is an effector-binding domain and contributes to the specificity of these receptors. Mutation in the LRR regions of NLRs can alter recognition specificity, which explains why LRR domains are highly polymorphic as a result of the strong co-evolutionary arms race between plants and pathogens (DeYoung and Innes, 2006; Ravensdale et al., 2011). Functional studies indicated that the N-terminal Toll/interleukin 1 receptor (TIR) domain in NLRs of flax or the coiled-coil (CC) domain in NLRs of wheat (*Sr33*, *Sr50*) mediate cell death signaling processes and initiate resistance responses via self-association (Bernoux et al., 2011; Casey et al., 2016). The molecular basis of rust race-specific resistance mechanisms has mostly been studied in flax rust (*Melampsora lini*) system due to the availability of genetic resources (Dodds et al., 2006).

For wheat, several race-specific resistance genes effective to stem rust have been cloned via map-based cloning combined with mutagenesis, which is time-consuming as it

requires extensive crosses and phenotypic analyses for positional fine mapping. The wheat stem rust resistance genes *Sr33, Sr35, Sr50*, *Sr21*, *Sr13*, *Sr60* were cloned using this strategy (Saintenac et al., 2013; Mago et al., 2015; Periyannan et al., 2017; Zhang et al., 2017; Chen et al., 2018, 2019). Most recently, rapid *R* gene cloning techniques have been developed such as MutRenseq, which combines mutagenesis and exome capture of NLR genes (Steuernagel et al., 2016), and AgRenSeq, which combines association genetics and *R* gene enrichment sequencing (Arora et al., 2019). MutRenseq is designed to detect sequence changes in NLRs in loss-of-resistance mutants derived from ethyl methanesulfonate (EMS) mutagenesis. This technique was utilized in cloning *Sr22* and *Sr45* in hexaploid wheat (Steuernagel et al., 2016). Alternatively, AgRenSeq enables rapidly cloning of *R* genes from any crop on the basis of pre-existing recombination events in natural populations and has been applied to clone two novel *Sr* genes in diploid wheat *Aegilops tauschii*, *Sr46* and *SrTA1662* (Arora et al., 2019).

Although race-specific resistance can provide almost complete immunity to certain pathotypes of the pathogen, the extensive use of monocultures comprised of a single wheat cultivar carrying one *R* gene promotes the rapid adaptation of rust pathogens (Ellis et al., 2014). This is explained by the selection pressure imposed on the pathogen to escape recognition by the *R* gene in the plant, which results in the evolution of new virulent rust races. In general, the lifetime of single *R* genes in the field is often very short, only lasting a few years after the release of resistant cultivars (Ellis et al., 2014).

Instead of utilizing single *R* genes, advances in the rapid isolation of *R* genes by techniques such as MutRenSeq and AgRenSeq allow for a strategy of stacking *R* genes and

transforming them as a gene cassette into the same elite line exploiting genetic engineering approaches (Steuernagel et al., 2016; Arora et al., 2019). This strategy of multi-*R* gene pyramids holds promise to provide more durable disease resistance as the pathogen needs to develop multiple single-step mutations to overcome pyramided *R* genes (Ayliffe et al., 2008; Ellis et al., 2014).

Adult plant resistance and its mechanisms

Adult plant resistance in disease breeding programs against rust fungi

Several wheat rust resistance genes confer non-race-specific resistance, which is effective against all races of a pathogen and even multiple pathogen species (Ellis et al., 2014). Adult plant resistance (*APR*) genes are usually in this class of resistance. *APR* genes mainly function at the adult stage and confer partial resistance, which is manifested by intermediate infection types with small pustules (Ellis et al., 2014). From an epidemiological perspective, the decreased degree of infection attributable to *APR* resistance is achieved by prolonging the pathogen's latency period as well as reducing infection frequency and spore production, and thus it is also referred to as “slow rusting” resistance (Niks et al., 2011). Interestingly, *APR* genes usually have pleiotropic phenotypes (Ellis et al., 2014). In addition to multi-pathogen resistance, most of these genes exhibit leaf tip necrosis in flag leaves (*Lr34*, *Lr67*) or pseudo black chaff (*Sr2*) (Singh, 1992). The biological reasons for those phenotypes are not understood since the mode of action of these genes remain elusive for the most part.

In contrast to race-specific *NLR*-mediated resistance, the quantitative *APR* genes have shown durable resistance in the field despite their extensive use in many modern wheat cultivars (Johnson, 1984; Ellis et al., 2014; Periyannan et al., 2017). Two of the most important *APR* genes, *Sr2*, *Lr34*, have been effective for over 100 years (Ellis et al., 2014). The deployment of *APR* genes in breeding programs is usually coupled with race-specific *R* genes to achieve the best immunity outcomes, because the "slow-rusting" *APR* genes alone are not sufficient to prevent disease completely (German and Kolmer, 1992; Vanegas et al., 2008; Milne et al., 2019).

Genetic and molecular mechanisms of adult plant resistance

Wheat *APR* genes *Lr34* and *Lr67* confer partial resistance against the three wheat rusts (*Pgt*, *Pst*, *Pt*) and powdery mildew (*Blumeria graminis* f. sp. *tritici*) in adult plants and are mostly effective in flag leaves (Ellis et al., 2014). Transgenic studies using *Lr34* and *Lr67* have shown these genes can also provide resistance against biotrophic or hemi-biotrophic pathogens in other cereals such as maize (Sucher et al., 2017), rice (Krattinger et al., 2016) and barley (Milne et al., 2019). These findings suggest that conserved molecular mechanisms underlying multi-pathogen resistance may exist, although it remains to be determined how these genes function to confer such a broad-spectrum resistance.

None of the cloned *APR* genes belongs to the *NLR* family. One of the cloned *APR* genes is *Yr36,* which provides durable and partial resistance to all races of the stripe rust

pathogen (*Pst*) under high temperature conditions (Fu et al., 2009). *Yr36* encodes a protein that contains a serine/threonine kinase domain and steroidogenic acute regulatory protein-related transfer (START) domain, both required for resistance (Fu et al., 2009). *Yr36* is proposed to play a role in protein phosphorylation in plant chloroplasts, which reduces peroxidase activity and increases H_2O_2 concentration, and therefore enhancing cell death and defense responses (Gou et al., 2015).

Another two cloned wheat *APR* genes are *Lr34* and *Lr67*, which are located on chromosomes 7D and 4D, respectively. *Lr34* and *Lr67* encode a putative ATP-binding cassette (ABC) transporter and a hexose transporter, respectively (Krattinger et al., 2009; Moore et al., 2015). Interestingly, the resistance alleles of *Lr34* (*Lr34res*) and *Lr67* (*Lr67res*) are each distinguished by just two-amino acid changes from their respective susceptibility alleles, and in both cases the resistance alleles arose by spontaneous mutations after wheat domestication (Krattinger et al., 2011, 2013; Moore et al., 2015). Recent research efforts have been directed at dissecting how these *APR* genes function. Krattinger and colleagues recently proposed that the substrate of the ABC transporter Lr34 is the phytohormone abscisic acid (ABA), and that Lr34-mediated ABA redistribution contributes to multi-pathogen resistance in several cereal crops consistent with a common pathway to confer resistance (Krattinger et al., 2019). Both Lr34res and Lr34sus proteins are functional ABA transporters during *in vitro* yeast assays. However, *in planta* assays showed that while both *Lr34* alleles were expressed at the mRNA transcript level, the Lr34sus protein was not detectable. Thus, these findings suggest that post-transcriptional or post-translational regulation of *Lr34* leads to modulation of ABA-mediated defense

responses (Krattinger et al., 2019). Study of wheat *Lr67* indicated that the susceptibility allele (*Lr67sus*) expresses a functional hexose transporter, while the resistance allele (*Lr67res*) has lost this function (Moore et al., 2015). However, *Lr67res* confers a dominant gain of resistance in transgenic wheat plants, and the deletion of this gene does not reproduce the resistance phenotype (Moore et al., 2015). Co-expression of *Lr67res* and *Lr67sus* in yeast suggests that *Lr67res* may have a dominant negative effect on glucose uptake (Moore et al., 2015), possibly mediated by the heterodimerization of Lr67res with the functional sugar transporters on A and B homeologs to reduce glucose uptake. Part of Chapter 3 addresses the role of an ortholog of *Lr67*, *Bradi1g69710*, in *Brachypodium distachyon* in defense responses to cereal rusts.

Non-host resistance and its mechanisms

The various definitions of non-host resistance

Non-host resistance (NHR) protects plants from the vast majority of the potential pathogenic threats, and thus it holds great potential to provide durable and broad-spectrum resistance (Heath, 2000). Microbes that fail to cause infection on a plant outside of its host range are called non-adapted pathogens for this particular plant species. Under strict definition, a non-host would be a plant species which is resistant against all isolates or races of a microbial pathogen (Mysore and Ryu, 2004; Bettgenhaeuser et al., 2014). However, this simple dichotomous definition of non-host status does not capture the various intermediate states that occur in plant-microbe interactions. For example, defining a plant

species as a non-host becomes ambiguous when only some accessions of this plant species are susceptible to the non-adapted pathogens or only certain isolates of the pathogen can actually infect this plant. Often, NHR phenotypes are quantitative and include a continuum of responses, from complete immunity, a range of intermediate interactions to almost complete susceptibility (Bettgenhaeuser et al., 2014; Omidvar et al., 2018). In the intermediate interaction phase, the pathogen may not complete its life cycle, which in the case of rust may mean formation of all infection structures but lack of sporulation (Bettgenhaeuser et al., 2014). I summarized different levels of NHR against rust fungi and their underlying mechanisms in Figure 1.2.

Studies of NHR may shed light on novel and durable genetic resistance resources as this type of resistance is non-race-specific and less likely to be overcome by pathogen evolution (Heath, 2000). Therefore, there has been a growing interest in improving our understanding of NHR against rust fungi and transferring that knowledge to agricultural host crops (Bettgenhaeuser et al., 2014).

Physical and molecular mechanisms associated with non-host resistance

Although there is considerable knowledge about Avr-R interactions in conferring resistance to adapted pathogens, key gaps in our understanding of molecular mechanisms in NHR need to be addressed (Mysore and Ryu, 2004). NHR consists of a continuum of plant defenses that include both physical and chemical barriers. The preformed and induced physical barriers include basic penetration incompatibility such as observed for cereal rusts

interacting with broad bean, and alternatively the bean rust pathogen *(Uromyces fabae*) with wheat (Zhang et al., 2011; Cheng et al., 2012). In these two pathosystems, few urediniospores were able to identify a stoma to produce an appressorium on the non-host plant. Callose deposition or papillae formation were also triggered to strengthen the plant cell wall at attempted infection sites (Figure 1.2). In addition to the physical barriers, induction of chemical defenses were also associated with these NHR responses such as phytoalexin production and expression of pathogenesis related (*PR*)-genes and genes linked to oxidative stress (Cheng et al., 2012).

NHR mechanisms in Arabidopsis had attracted researchers' attention due to the considerable resources available for this model plant. Several studies suggest NHR can occur at both pre-haustorial or post-haustoria phases in the interactions of Arabidopsis with its non-adapted pathogens, such as *Blumeria graminis* f. sp. *hordei* (*Bgh*, the causal agent of barley powdery mildew) and *Puccinia triticina* (*Pt*, the causal agent of wheat leaf rust) (Lipka et al., 2005; Ellis, 2006; Bettgenhaeuser et al., 2014). Only about 10% and 0.2% of spores of *Bgh* and *Pt*, respectively, were capable of forming haustoria in Arabidopsis plants, indicating a significant role of pre-haustoria NHR defenses in these interactions (Collins et al., 2003; Shafiei et al., 2007). Three NHR genes, *PEN1* (PENETRATION), *PEN2*, *PEN3*, which are functional at the pre-haustorial stage in Arabidopsis against *Bgh* have been cloned. Single mutations in these genes led to increased formation of haustoria of mildew pathogens, but no colonies of viable spores were produced (Collins et al., 2003; Ellis, 2006; Stein et al., 2006). *PEN1* encodes a membrane-associated syntaxin, which is involved in vesicle-mediated exocytosis at the penetration site that contributes to papillae accumulation

and restricts fungal growth. *PEN2* and *PEN3* encode a glycosyl hydrolase and an ABC transporter, respectively, and may act together in a pathway to deliver unknown antifungal metabolites to the pathogen invasion site (Collins et al., 2003; Lipka et al., 2005; Stein et al., 2006). Therefore, in pre-haustorial resistance, organelle and vesicle movement could mediate a battery of NHR-associated defense responses, and these activities could be activated by a PAMP such as fungal chitin or chemical signals secreted during pathogen entry (Mysore and Ryu, 2004; Ellis, 2006).

In the case of the pathogen escaping pre-haustorial NHR, infection appears to be stopped by SA-dependent HR responses (Stein et al., 2006). Mutations of both pre- and post-haustorial resistance genes led to enhanced susceptibility of Arabidopsis to the non-adapted pathogen, *Erysiphe pisi*, the causal agent of pea mildew (Stein et al., 2006). It is unknown if *R* gene-mediated HR responses can occur during the post-haustorial phase, as the localized cell death induced by non-adapted pathogens could be attributed to the recognition of a pathogen effector or by an amplified signaling flux after detection of PAMPs (Ellis, 2006).

NHR research in the Arabidopsis-*Pt* pathosystem has also suggested the reduced number of haustoria could be explained by papillae formation and thickening cell wall during the rust penetration stage (Yun et al., 2003; Shafiei et al., 2007; Loehrer et al., 2008). Those haustoria that did form were aberrantly shaped (Shafiei et al., 2007). In addition, callose deposition, antimicrobial phenolic compounds and salicylic acid (SA)-dependent defense pathways may be key components affecting penetration and haustoria formation in this pathosystem (Mellersh and Heath, 2003). Other studies in Arabidopsis showed that

NHR may involve the deployment of different defense responses against different non-adapted pathogens. For instance, *Pt* induced both jasmonic acid (JA) and salicylic acid (SA) signaling pathways as well as the production of reactive oxygen species, while *Bgh* did not induce JA signaling in Arabidopsis (Shafiei et al., 2007; Loehrer et al., 2008). Briefly, research of NHR in Arabidopsis discovered that most rust colonization was arrested at the production of haustorial mother cells, with successful development of haustorium in only rare cases (Bettgenhaeuser et al., 2014). In these cases, localized plant cell death to rust fungi was also common in Arabidopsis, and these reactions are polygenic and active responses associated with SA signaling pathway (Ayliffe et al., 2010; Zellerhoff et al., 2010) (Figure 1.2).

It remains as a question how we can utilize NHR components or genes to enhance disease resistance in a host crop of interest. NHR mechanisms identified in Arabidopsis against rust fungi may be difficult to translate to provide disease protection in cereals because of the distant relationship between monocots and dicots, considering that these groups split ~ 200 million years ago (Mya) (Schulze-Lefert and Panstruga, 2011). Ayliffe and co-authors explored NHR in a rice-rust pathosystem given that rice is a non-host to all rust species (Ayliffe et al., 2010, 2011). Rice interactions with cereal rust species revealed that rice cultivars display natural variation in levels of resistance to *Pgt* and *Pst*, and in contrast to the interactions between rust and Arabidopsis and bean, cereal rust fungi established haustoria in mesophyll cells and formed small, non-sporulating colonies (Ayliffe et al., 2010, 2011). Different rice cultivars exhibited subtle distinctions in resistance phenotypes to rust pathogens, and this study suggest that multi-layered defense

responses are triggered in the rice-rust pathosystem, including callose deposition and the accumulation of reactive oxygen species (Ayliffe et al., 2010, 2011) (Figure 1.2). Unfortunately, phenotypes associated with NHR in rice in response to rust can be affected by environmental conditions (Ayliffe et al., 2011, 2013a), which complicates the tractability of the system. These challenges and the subtle differences in phenotypes have prevented scientists from using rice as a system to understand NHR mechanisms against rust fungi.

Non-host resistance from an evolutionary perspective

In general, a continuum of infection outcomes can be detected when examining NHR (Bettgenhaeuser et al., 2014). Both PTI and ETI contribute to NHR, but the relative contribution is determined by the phylogenetic divergence between the host and non-host species (Figure 1.2). It seems that the relative contribution of PTI to NHR increases with more divergence between the host and non-host species. Otherwise, ETI likely plays a more critical role in the reactions if the host and non-host species are evolutionary related, as the non-host may retain some conserved NLRs that recognize effectors of the non-adapted pathogen due to past evolutionary interactions. According to this idea, the underlying mechanisms of NHR could be explained by the lack of co-evolution of pathogen effectors and plant immune receptors in distant species, which then leads to disease incompatibility (Schulze-Lefert and Panstruga, 2011).

Applications of non-host resistance for enhancing disease resistance

Several cases illustrate the potential applications and translation of non-host resistance to crops (Wulff et al., 2011). Transfer of a PTI component from a non-host to a host can increase disease resistance. An example of this is given by the success that transferring the leucine rich repeat receptor-kinase EFR, a PRR, from Arabidopsis to tomato and *Nicotiana benthaminana* conferred broad-spectrum resistance against several bacterial pathogens (Lacombe et al., 2010). The transfer of an ETI component also holds promise in providing genetic resistance. For example, rice plants transferred with a *NLR* gene from maize exhibited strong resistance (hypersensitive response, HR) against rice leaf streak pathogen *Xanthomonas oryzae* pv. *oryzicola* (Zhao et al., 2005). Similarly, soybean containing a *NLR* gene that was derived from a close diploid relative of soybean, pigeonpea, showed full resistance against Asian soybean rust (ASR). In this study, 52 accessions of pigeonpea displayed a range of disease phenotypes against ASR (Kawashima et al., 2016). This is a good example of mining genetic resources of disease resistance in the close relatives of the major crops. However, it is questionable if transferring a single gene like *NLR* can confer durable disease resistance despite that it is derived from a non-host plant (Bettgenhaeuser et al., 2014). In fact, the deployment of a cultivar introgressed with one single non-host resistance gene may also enforce strong selection pressure for the evolution of the non-adapted pathogen, and thus leading to potential erosion of NHR. In return, pathogen adaptation may lead to the accumulation of new virulence genes and eventually give rise to the development of a pathogen able to infect the gene donor species which was originally not a host (Bettgenhaeuser et al., 2014). A concerning case is the gain of

virulence in *Puccinia striiformis* f. sp. *tritici* to rye after the introgression of a rye resistance gene, *Yr9*, into wheat (Wellings, 2012a; Bettgenhaeuser et al., 2014). Therefore, the significance of the genetic erosion of NHR causing plant vulnerability is specifically affected by the agronomic values of the gene donor species (Bettgenhaeuser et al., 2014). With this in mind, the heterologous transfer of genes from a non-host plant like *Brachypodium distachyon* to wheat may be a better option since *B. distachyon* is not agronomically significant (Bettgenhaeuser et al., 2014).

Chapter 3 in this dissertation utilizes a reverse genetics approach to investigate the genetic factors in *B. distachyon* that are associated with disease defense responses associated with NHR against rust fungi, specifically *Pgt*.

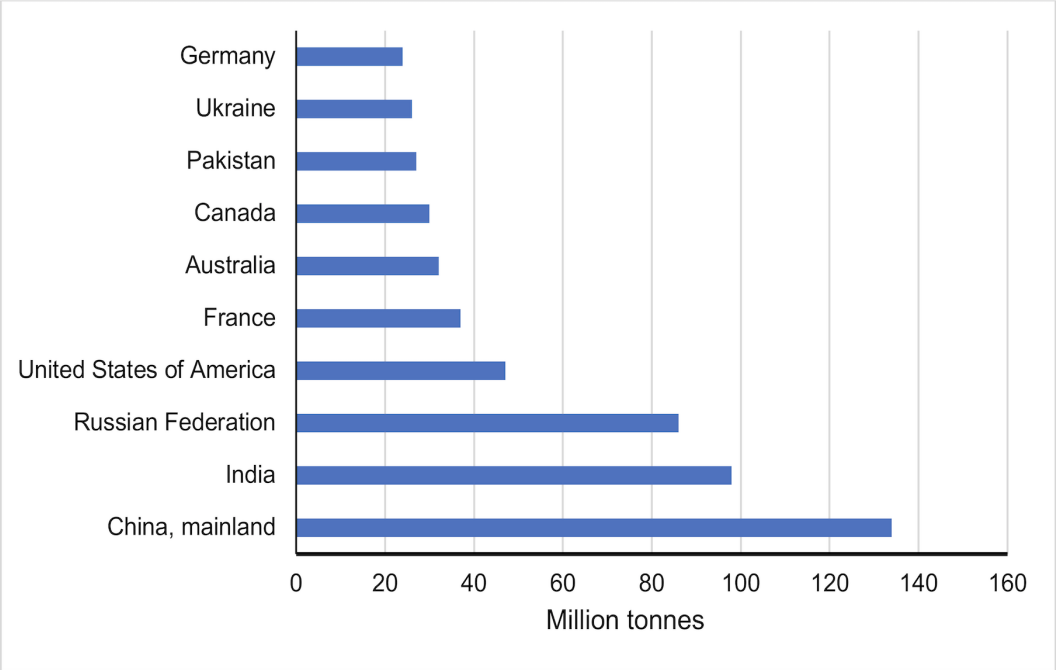


Figure 1.1. Top ten producers of wheat in 2017 (FAO, 2017).
Data accessed on February 14, 2019.

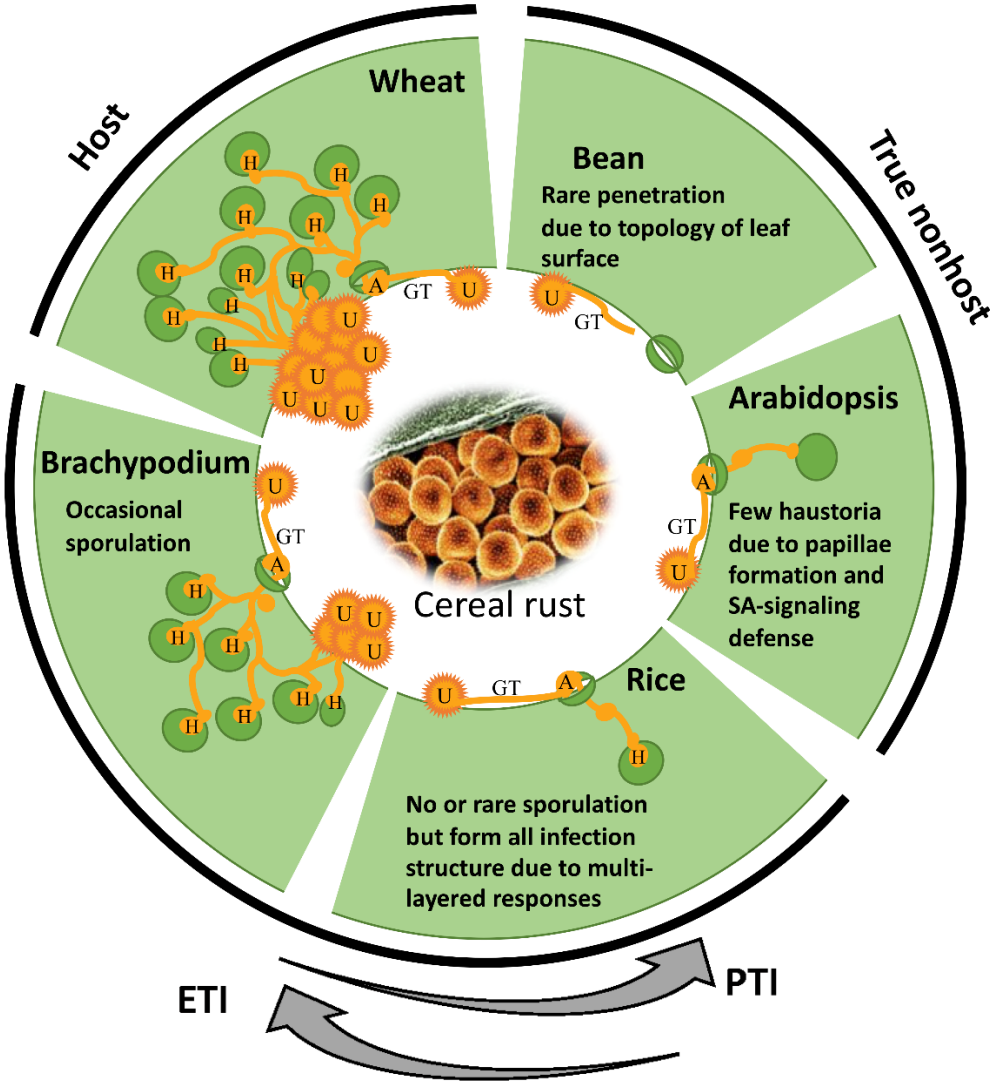


Figure 1.2. An overview of non-host resistance (NHR) mechanisms in the interactions of cereal rust fungi and different host.
PTI: PAMP-triggered immunity. ETI: Effector-triggered immunity. The arrows on the bottom picture represent the relative contribution of PTI and ETI to NHR. Diagrams shown in orange represent fungal structure of cereal rust fungus, including urediniospore (U), germ tube (GT), appressorium (A), haustorium (H). The green oval indicates the mesophyll cell of the plant while the two half-crescent structure in green shows an opening stoma.

Chapter 2. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation[1]

Contributions to Chapter 2

In this Chapter, I was responsible for the following parts:

- Extraction of high molecular weight DNA from *Puccinia graminis* f. sp. *tritici* (*Pgt*) isolate Ug99 at the Biosafety Level 3 Laboratory at the University of Minnesota. Submission of samples for PacBio sequencing at Leidos Biomedical Research (Supplemental Table 2).
- Extraction of genomic DNA for Illumina sequencing of Ug99 and four other isolates of the Ug99 lineage at the Biosafety Level 3 Laboratory at the University of Minnesota. Submission of samples to the University of Minnesota Genomic Center (Supplemental Table 3).

[1] This work was accepted in *Nature Communications*. **Li, F.,** Upadhyaya, N., Sperschneider, J., Matny, O., Nguyen-Phuc, H., Mago, R., Raley, C., Miller, M., Silverstein, K., Henningsen, E., Hirsch, C., Visser, B., Pretorius, Z., Steffenson, B., Schwessinger, B., Dodds, P.N., Figueroa, M., 2019. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *bioRxiv* 692640. *In Press*. *Nature Communications*. https://doi.org/10.1038/s41467-019-12927-7

Authors Contributions: M.F. and P.N.D. conceptualized the project, acquired funding and supervised the work. B.V. and Z.A.P. provided study materials. F.L., N.M.U., C.R., O.M., B.S., R.M., and B.J.S. acquired experimental data. F.L., N.M.U., J.S., B.S., B.J.S., H.N.P., P.N.D, K.S., E.H., M.E.M., and C.D.H. conducted data analysis. M.F., P.N.D. and F.L. drafted the manuscript. All authors contributed to review and editing.

- *De novo* genome assembly of Ug99 using PacBio-based long reads, including testing two independent long read assemblers, optimizing genome assembly polishing steps and quality assessment (Supplemental Table 4).
- Haplotype phasing and bin assignment of contigs from Ug99; telomere detection for the Ug99 isolate (Figure 2.1, Supplemental Table 7).
- Manual curation of chimeric contigs in Ug99 (Supplemental Table 8).
- Haplotype assignment of Ug99 contigs using a read subtraction mapping approach (Figure 2.3b-d, Supplemental Table 7).
- Haplotype comparisons at bin and entire haplotype levels between Ug99 and *Pgt*21-0 (Table 2.1, Figure 2.2, 2.4, Supplemental Table 6, 9, Supplemental Figure 2, 3a, 3c-e, 4a-c).
- Detection of structural variation of 57 kbp insert in *AvrSr35* allele of Ug99 (Supplemental Figure 1).
- Structural variation analysis of haplotypes within and among *Pgt* isolates (Table 2.1, Supplemental Table 6, 9, 15, Supplemental Figure 8).
- Gene prediction and functional annotation for Ug99 and *Pgt*21-0 (Supplemental Table 13).
- Comparison of mating type loci between the Ug99 and *Pgt*21-0 (Supplemental Figure 6).
- Management of submissions of both genome assemblies to public repositories.

Parasexuality contributes to diversity and adaptive evolution of haploid (monokaryotic) fungi. However, non-sexual genetic exchange mechanisms are not defined in dikaryotic fungi (containing two distinct haploid nuclei). Newly emerged strains of the wheat stem rust pathogen, *Puccinia graminis* f. sp. *tritici* (*Pgt*), such as Ug99, are a major threat to global food security. Here we show that Ug99 arose by somatic hybridisation and nuclear exchange between dikaryons. Fully haplotype-resolved genome assembly and DNA proximity analysis reveal that Ug99 shares one haploid nucleus genotype with a much older African lineage of *Pgt*, with no recombination or chromosome reassortment. These findings demonstrate that nuclear exchange between dikaryotes is an important factor in generating genetic diversity and can facilitate the emergence of new lineages in asexual fungal populations.

Introduction

Generation of genetic diversity is crucial for the evolution of new traits, with mutation and sexual recombination as the main drivers of diversity in most eukaryotes. However, many species in the fungal kingdom can propagate asexually for extended periods and therefore understanding alternative mechanisms contributing to genetic diversity in asexual populations has been of great interest (Stajich et al., 2009; Fleißner and Serrano, 2016). Some fungi can use a parasexual mechanism to exchange genetic material independently of meiosis (Fleißner and Serrano, 2016). This process involves anastomosis of haploid hyphae and fusion of two nuclei to generate a single diploid nucleus, which subsequently undergoes progressive chromosome loss to generate recombinant haploid

offspring. Parasexuality has been described in members of the ascomycete phylum (64% of described fungal species) in which the dominant asexually propagating form is haploid (Spatafora et al., 2017). However, in basidiomycete fungi (34% of described species), the predominant life stage is generally dikaryotic, with two different haploid nuclei maintained within each individual (Spatafora et al., 2017). The role of non-sexual genetic exchange between such dikaryons in generating genetic diversity is not known.

Basidiomycetes include many fungi with critical ecosystem functions, such as wood decay and plant symbiosis, as well as agents of important human and plant diseases (Stajich et al., 2009). Rust fungi (subphylum Pucciniomycotina) comprise over 8,000 species including many pathogens of major agricultural and ecological significance (Aime et al., 2017). These organisms are obligate parasites with complex life cycles that can include indefinite asexual reproduction through infectious dikaryotic urediniospores. Early researchers speculated that rust fungi can exchange genetic material during the asexual phase based on the isolation of new strains, after co-infection with two potential parental isolates, with novel virulence phenotypes (Watson, 1957; Ellingboe, 1961; Flor, 1964; Bartos et al., 1969). However, these hypotheses could not be tested molecularly at the time. Some naturally occurring rust pathotypes have been suggested to have arisen by somatic hybridisation and genetic exchange based on limited molecular evidence of shared isozyme or random amplified polymorphic DNA (RAPD) markers (Burdon et al., 1981; Park et al., 1999). Mechanisms underlying genetic exchange are unknown, but may involve hyphal anastomosis followed by nuclear exchange and/or nuclear fusion and recombination (Park and Wellings, 2012). Recent advances in assembling complete karyon sequences in rust

fungi (Miller et al., 2018; Schwessinger et al., 2018) provide the opportunity to definitively detect and discriminate between nuclear exchange and recombination.

The Ug99 strain (race TTKSK) of the wheat stem rust pathogen *Puccinia graminis* f. sp. *tritici* (*Pgt*) presents a significant threat to global wheat production (Singh et al., 2015). It was first detected in Uganda in 1998 and described in 1999 (Pretorius et al., 2000), and has since given rise to an asexual lineage that has spread through Africa and the Middle East causing devastating epidemics (Singh et al., 2015). The origin of the Ug99 lineage is unknown, although it is genetically distinct from other *Pgt* races (Visser et al., 2011; Olivera et al., 2015). To resolve the genetic makeup of Ug99, we generated a haplotype-phased genome reference for the original Ug99 isolate collected in Uganda (Pretorius et al., 2000). In addition, we also generated a similar reference for an Australian *Pgt* isolate of pathotype 21-0 (Park, 2007; Upadhyaya et al., 2015). This isolate is a member of a longstanding asexual lineage that has been predominant in southern Africa since the 1920's and spread to Australia in the 1950's (Park, 2007; Pretorius et al., 2007; Visser et al., 2019). Our study shows that Ug99 shares one haploid nucleus genotype with *Pgt*21-0, with no recombination or chromosome reassortment. This indicates that Ug99 arose by somatic hybridisation and nuclear exchange between an African member of the *Pgt* 21 lineage and an unknown isolate. Thus, nuclear exchanges between dikaryotic fungi can contribute to the emergence of new variants with significant epidemiological impacts.

Materials and Methods

Fungal stocks and plant inoculation procedures

Pgt isolates Ug99 (Pretorius et al., 2000) collected in Uganda, UVPgt55, UVPgt59, UVPgt60 and UVPgt61 collected in South Africa (Visser et al., 2011; Pretorius et al., 2012) were transferred to the Biosafety Level 3 (BSL-3) containment facility at the University of Minnesota for growth and manipulation. Samples were purified by single pustule isolation and then amplified by 2-3 rounds of on the susceptible wheat cultivar McNair. Virulence pathotypes and purity of each isolate were confirmed by inoculation onto the standard wheat differential set (Supplementary Table 1). *Pgt*21-0 was first isolated in Australia in 1954 and has been previously described (Park, 2007; Upadhyaya et al., 2015). North American isolate CRL 75-36-700-3 (pathotype SCCL) and Kenyan isolate 04KEN156/04 (pathotype TTKSK) were described previously (Olivera et al., 2015; Aime et al., 2017). For rust inoculations, urediniospores retrieved from -80°C were heat treated (45°C for 15 min) and suspended in mineral oil (Soltrol 170, Philips Petroleum, Borger, TX, U.S.A.) at 14 mg/ml. Seven-day-old seedlings were spray-inoculated at 50 µl/plant and kept in a dark mist chamber at 22–25°C and 100% humidity for 16 h. Subsequently, plants were exposed to light (150–250 µmol photons s^{-1} m^{-2}) for 5 h in the mist chamber and then transferred to a controlled growth chamber (18 h/6 h of light/dark, 24°C/18°C for day/night, 50% relative humidity). Spores were collected 9 and 14 days post inoculation (dpi) and maintained at -80°C.

DNA extraction and sequencing of rust isolates

High molecular weight DNA of Ug99 and *Pgt*21-0 was extracted from 300-350 mg urediniospores as described (Schwessinger and Rathjen, 2017), with the following modifications: 1) for Phenol:Chloroform:Isoamyl alcohol extractions, samples were centrifuged at 4°C and 5,000 x g for 20 mins; 2) a wide-bore 1mL pipette tip was used to transfer the DNA pellet; 3) samples were incubated for 1 h at 28°C with 200-250 rpm shaking to dissolve the final DNA pellet. Double stranded DNA concentration was quantified using a broad-range assay in a Qubit Fluorometer (Invitrogen, Carlsbad, CA, U.S.A.) and a NanoDrop (Thermo Fisher Scientific, Waltham, MA, U.S.A.). Approximately 10 µg DNA from Ug99 and *Pgt*21-0 was sequenced using PacBio SMRT sequencing (Pacific Bioscience, Menlo Park, CA, U.S.A.) at either the Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc. (Frederick, MD, U.S.A.) or the Ramaciotti Centre (Sydney, Australia), respectively. DNA was concentrated and cleaned using AMPure PB beads for Ug99 or AMPure XP beads for *Pgt*21-0 (Pacific Biosciences). DNA quantification and size assessment were conducted using a NanoDrop (Thermo Fisher Scientific) and 2200 TapeStation instruments (Agilent Technologies, Santa Clara, CA, U.S.A.). DNA was sheared to a targeted average size of 20 kb using G-tubes (Covaris, Woburn, MA, U.S.A.). Libraries were constructed following the 20 kb Template Preparation BluePippin Size-Selection System protocol (Pacific Biosciences) using a BluePippin instrument (Sage Science, Beverly, MA, U.S.A.) with a 0.75% agarose cassette and a lower cutoff of 15 kbp. For Ug99, 5 SMRT cells were sequenced on a PacBio Sequel platform using P6-C4 chemistry, the Sequel Binding Kit 2.0 (Pacific Biosciences),

diffusion loading, 10-hour movie lengths and Magbead loading at 2 pM (3 cells) or 4 pM (2 cells). In addition, 4 SMRT cells were run on PacBio RSII sequencer using P6-C4 chemistry, with 0.15 nM MagBead loading and 360-min movie lengths. For *Pgt*21-0, 17 SMRT cells were run on the RSII platform using P6-C4 chemistry, Magbead loading (0.12-0.18 nM) and 240-min movie lengths.

Genomic DNA for Illumina sequencing was extracted from 10-20 mg urediniospores of Ug99, UVPgt55, 59, 60 and 61 using the OmniPrep™ kit (G-Biosciences, St. Louis, MO, U.S.A.) following the manufacturer's instructions. TruSeq Nano DNA libraries were prepared from 300 ng of DNA, and 150bp paired-end sequence reads were generated at the University of Minnesota Genomics Center on the Illumina NextSeq 550 platform using Illumina Real-Time Analysis software version 1.18.64 for quality-scored base calling.

De novo **long read assembly**

Genome assemblies of Ug99 and *Pgt*21-0 were built from PacBio reads using Canu version 1.6 (Koren et al., 2017) with default parameters and an estimated genome size of 170 Mbp. Assemblies were polished with the Arrow algorithm using the raw PacBio reads in the sa3_ds_resequencing pipeline in pbsmrtpipe workflow within SMRTLINK/5.1.0 (Pacific BioSciences). Assemblies were further polished by two rounds of Pilon (Walker et al., 2014) with the option fix --all using Illumina reads from Ug99 (this work) or *Pgt*21-0 (NCBI SRA run Accession# SRR6242031). A BLASTN search (version 2.7.1) against

the NCBI nr/nt database (downloaded on 4/11/2018) with E-value set as $1e^{-10}$ identified two contigs in the Ug99 assembly with significant hits to plant rRNA and chloroplast sequences and these were removed.

PacBio and Illumina reads were mapped to the assembly using BWA-MEM (version 0.7.17) (Li and Durbin, 2009) and BAM files were indexed and sorted using SAMtools (version 1.9) (Li et al., 2009). Read coverage analysis using genomeCoverageBed in BEDtools (version 2.27.1) (Quinlan and Hall, 2010) identified 144 small contigs (< 50 kbp) in the Ug99 assembly with low coverage (< 2X) for both short and long reads and these contigs were also excluded from the final assembly. Genome assembly metrics were assayed using QUAST (version 4.3) (Gurevich et al., 2013). Genome completeness was assessed via benchmarking universal single-copy orthologs (BUSCOs) of the basidiomycota as fungal lineage and *Ustilago maydis* as the reference species for AUGUSTUS gene prediction (Stanke and Morgenstern, 2005) in the software BUSCO v2.0 (genome mode) (Simão et al., 2015). Telomeric sequences were identified using either a high stringency BLAST with 32 repeats of TTAGGG as query or a custom python script to detect at least five CCCTAA or TTAGGG repeats in the assemblies. Repeats of at least 60 bp length and occurring within 100 bp of the contig end were defined as telomeric sequences.

Detection of alternate contigs and bin assignment

To identify contigs representing corresponding haplotypes (Figure 2.1) 22,484 predicted *Pgt* gene coding sequences (Upadhyaya et al., 2015) were screened against the genome assemblies using BLITZ (Blat-like local alignment) in the Biokanga Tool set, (https://github.com/csiro-crop-informatics/biokanga/releases/tag/v4.3.9). For each gene the two best hits (likely alleles) in the assembly were recorded. Contigs sharing best hits for at least five genes were selected as potential haplotype pairs and their sequence collinearity was examined by alignment and similarity plotting using D-genies (Cabanettes and Klopp, 2018). Contigs representing contiguous or syntenous haplotypes were grouped together as bins.

Validation of a 57 kbp-insert in *AvrSr35*

Contigs containing the *AvrSr50* and *AvrSr35* gene sequences were identified by BLASTN search against customized databases for the Ug99 and *Pgt*21-0 genome assemblies. Illumina and PacBio reads of Ug99 mapped to the genome assembly were visualized in the Integrative Genomics Viewer (IGV). To validate the presence of the 57 kbp-insert in *AvrSr35*, flanking and internal sequences were amplified from genomic DNA extracted using the OmniPrep™ kit (G-Biosciences) from urediniospores of Ug99, 04KEN156/04, and CRL 75-36-700-3. PCR was performed using Phusion high-fidelity DNA polymerase according to the manufacturer's recommendations (New England BioLabs Inc., Ipswich, MA, U.S.A.) and primers listed in Supplementary Table 15. The

amplified PCR products were separated by electrophoresis on a 1% agarose gel and stained using SYBR Safe DNA gel stain (Invitrogen). Specific bands were cleaned using NucleoSpin gel clean-up kit (Takara Bio, Mountain View, CA, U.S.A.) for subsequent Sanger sequencing and alignment to *AvrSr35* alleles. Gene models in the *AvrSr35* and *AvrSr50* locus were depicted using GenomicFeatures (Lawrence et al., 2013) and ggbio (Yin et al., 2012a) in a custom R script.

Haplotype assignment by read cross-mapping and subtraction

Illumina reads from *Pgt*21-0 (NCBI SRR6242031) were trimmed ("Trim sequences" quality limit = 0.01) and mapped to the Ug99 reference assembly using the "map reads to reference" tool in CLC Genomics Workbench version 10.0.1 or later with high stringency parameters (similarity fraction 0.99, length fraction 0.98, global alignment). Unmapped reads (Ug99-subtracted reads) were retained and then mapped back to the *Pgt*21-0 assembly contigs using the same parameters. The original *Pgt*21-0 reads were also mapped to the *Pgt*21-0 assembly and the read coverage for each contig compared to the Ug99-subtracted reads. Contigs with very low coverage (<2X total and <10% of the original read coverage) with the Ug99-subtracted reads were designated as karyon A (Figure 2.3, Supplementary Table 7). Contigs with substantial coverage of Ug99-subtracted reads (>10% of the original read coverage) were designated as karyon B. Contigs with ambiguous read mapping data, including those with low coverage in the original unsubtracted reads or covered by largely non-uniquely mapping reads were left as unassigned. Read mapping to all contigs was confirmed by visual inspection of coverage graphs and read alignments in

the CLC Genomics Workbench browser. Potential chimeric contigs were identified as containing distinct regions with either high or no coverage with the Ug99-subtracted reads (Supplementary Figure 4). For subsequent comparison and analyses, these contigs were manually split into their component fragments which were designated as haplotype A or B accordingly (Supplementary Table 8). The same process was followed in reverse for the assignment of the A and C haplotype contigs in Ug99.

Sequence comparisons of genome assemblies

Haplotype sequences of the *AvrSr50*/*AvrSr35* chromosome as well as the full haploid genomes were aligned using MUMmer4.x (Kurtz et al., 2004) with nucmer -maxmatch and other parameters set as default and the alignment metrics summarized in MUMmer dnadiff. Structural variation between haplotypes was determined using Assemblytics (Kurtz et al., 2004) from the MUMmer delta file with a minimum variant size of 50 bp, a maximum variant size of 100 kbp, and a unique sequence length for anchor filtering of 10 kbp. Haplotype dot plot alignments were generated using D-genies (Cabanettes and Klopp, 2018).

Read coverage analysis and SNP calling on haplotypes

Illumina reads from Ug99 and *Pgt*21-0 were each mapped against the Ug99 and *Pgt*21-0 assemblies in CLC Genomics Workbench (similarity fraction 0.98, length fraction

0.95). For each assembly the mean coverage per base was calculated per 1,000 bp interval ("window") using samtools bedcov. Read coverage frequency normalized to the mean coverage of each haplotype was graphed as a violin plot using seaborn 0.9.0 package (https://seaborn.pydata.org/) using a custom python script. To detect SNPs between two haplotypes, Illumina read pairs of *Pgt*21-0 that mapped uniquely to either the *Pgt*21-0 A or B haplotype contigs were extracted. Similarly, Ug99-derived read pairs that uniquely mapped to either the A or C haplotype contigs of Ug99 were extracted. These read sets were then separately mapped to the two assemblies in CLC Genomics Workbench (similarity fraction 0.99, length fraction 0.98). Variant calling was performed using FreeBayes v.1.1.0 (Garrison and Marth, 2012) with default parameters in parallel operation and SNPs were filtered using vcffilter of VCFlib (v1.0.0-rc1, https://github.com/vcflib/vcflib) with the parameter -f "QUAL > 20 & QUAL / AO > 10 & SAF > 0 & SAR > 0 & RPR > 1 & RPL > 1". Homozygous and heterozygous SNPs were extracted by vcffilter -f "AC > 0 & AC = 2" and -f "AC > 0 & AC = 1", respectively. SNP statistics were calculated using vcfstats of VCFlib.

Hi-C data analysis and scaffolding

A Hi-C library was constructed with the ProxiMeta Hi-C kit from Phase Genomics v1.0 containing the enzyme *Sau*3A from ~150 mg of dried urediniospores of *Pgt*21-0 following the standard protocol with minor modifications. Spores were washed in 1 mL 1X TBS buffer twice before cross-linking. After quenching of the crosslinking reaction, all liquid was removed and spores were frozen in liquid nitrogen. Spores were then lysed using

cryogenic bead beating with two 5 mm steel beads shaking twice for 45 sec at 25 Hz using TissueLyser II (Qiagen). Lysis buffer was added and spores vortexed until full suspension. Reverse cross-linking was performed at 65°C with 700 rpm horizontal shaking for 18 h. The Hi-C library was sequenced (150 bp paired-end reads) on the NextSeq 550 System using the Mid-Output Kit at the Ramaciotti Centre. The raw Hi-C reads were processed with the HiCUP pipeline version 0.7.1 (Wingett et al., 2015) (maximum di-tag length 700, minimum di-tag length 100, --re1 ^GATC, Sau3A), using bowtie2 as the aligner (Langmead and Salzberg, 2012) and the *Pgt*21-0 genome assembly as the reference. SAM files of the filtered di-tags were parsed to extract cis-far pairs (pairs located on the same contig and >10 kbp apart) and trans pairs (located on different contigs). The numbers of trans pairs connecting each pair of contigs was extracted from this data. We compared the number of Hi-C read pair connections between contigs of the same or different haplotypes, either within bins, within chromosomes or at the whole genome level (all haplotype-assigned contigs). We used a X^2 test to assess the deviation of each read pair distribution from a 1:1 ratio.

For scaffolding, the raw Hi-C reads were first mapped to the *Pgt*21-0 assembly using BWA-MEM (Li et al., 2009) and processed using the Arima Genomics pipeline (https://github.com/ArimaGenomics/mapping_pipeline/blob/master/01_mapping_arima.sh). Scaffolding was performed using SALSA 2.2 (Ghurye et al., 2017) on the full set of contigs, as well as independently on the haplotype A or B sets of contigs (including unassigned contigs). Invalid scaffold linkages between adjacent telomeres, which occur as an artefact of telomere co-location within the nucleus, were discarded. The three sets of

scaffolds were compared with the bin and haplotype assignment information to find overlaps and the resulting chromosome pseudomolecules were constructed by concatenating ordered contigs with 100 Ns inserted between contigs. To confirm translocations detected in the A and B chromosome sets, contigs that spanned the translocation site were separated into two fragments at the junction point and the SALSA scaffolding process was repeated on the full genome contig assembly. To detect nucleus-specific cross-links between chromosomes, HiCUP analysis was performed using the chromosome pseudomolecules as the reference assembly and the proportion of trans linkages between chromosomes of the same or different haplotype computed.

Gene prediction and functional annotation

The genome assemblies of Ug99 and *Pgt*21-0 (as chromosome pseudomolecules for *Pgt*21-0) were annotated using the Funannotate pipeline (https://github.com/nextgenusfs/funannotate). Contigs were sorted by length (longest to shortest) and repetitive elements were soft-masked using RepeatModeler (v1.0.11) and RepeatMasker (v4.0.5; http://www.repeatmasker.org/) with RepBase library (v. 23.09) (Bao et al., 2015). RNAseq libraries from *Pgt*21-0 (Supplementary Table 16) (Upadhyaya et al., 2015; Chen et al., 2017) were used for training gene models. In the training step, RNA-seq data were aligned to the genome assembly with HISAT2 (Kim et al., 2015). Transcripts were reconstructed with Stringtie (v1.3.4d) (Pertea et al., 2015). Genome-guided Trinity assembly (v2.4.0) (Grabherr et al., 2011) and PASA assembly (v2.3.3) (Haas et al., 2008) were performed. To assist in predicting effector-like genes, stringtie-

aligned transcripts were used in CodingQuarry Pathogen Mode (v2.0) (Testa et al., 2015). The prediction step of funannotate pipeline (funannotate predict) was run with --ploidy 2, --busco_db basidiomycota and default parameters. Transcript evidence included Trinity transcripts, Pucciniamycotina EST clusters downloaded from the JGI MycoCosm website (http://genome.jgi.doe.gov/pucciniomycotina/pucciniomycotina.info.html, April 24, 2017), and predicted transcript sequences of haustorial secreted proteins (Upadhyaya et al., 2015). Transcript evidence was aligned to the genome using minimap2 v2.1.0 (Li, 2018) and the protein evidence was aligned to genome via Diamond (v0.9.13)/Exonerate (v2.4.0) (Slater and Birney, 2005) using the default UniProtKb/SwissProt curated protein database from funannotate. *Ab initio* gene predictor AUGUSTUS v3.2.3 (Stanke and Morgenstern, 2005) was trained using PASA data and GeneMark-ES v4.32 (Lomsadze et al., 2005) was self-trained using the genome assembly. Evidence Modeler was used to combine all the above evidence using default weight settings except that the weight of PASA and CodingQuarry_PM were both set to 20. tRNA genes were predicted using tRNAscan-SE v1.3.1 (Lowe and Chan, 2016). Gene models including UTRs and alternative spliced transcripts were updated using RNAseq data based on Annotation Comparisons and Annotation Updates in PASA. Funannotate fix was run to validate gene models and NCBI submission requirements. Genome annotation was assessed using BUSCO v2.0 (transcript and protein modes) (Simão et al., 2015). For functional annotation protein coding gene models were firstly parsed using InterProScan5 (v5.23-62.0) to identify InterPro terms, GO ontology and fungal transcription factors (Jones et al., 2014). Pfam domains were identified using PFAM v. 32.0, and carbohydrate hydrolyzing enzymatic domains (CAZYmes) were annotated using dbCAN v7.0 (Yin et al., 2012b). Diamond blastP (Buchfink et al., 2015)

was used to search UniProt DB v. 2018_11 (Consortium, 2018) and MEROPS v. 12.0 (Rawlings, 2016) databases to aid in functional annotation. BUSCO groups were annotated with Basidiomycota models and eggNOG terms were identified using eggNOG-mapper v1.0.3 (Huerta-Cepas et al., 2015). Putative mating-type loci in *Pgt*21-0 and Ug99 were identified by BLAST search with the pheromone peptide encoding genes (*mfa2* or *mfa3*) and pheromone mating factor receptors (*STE3.2* and *STE3.3*) from the *a* locus, and *bW*/*bE* transcription factors from the *b* locus previously identified in *Pgt* (Cuomo et al., 2017). Protein sequences were aligned in Clustal Omega (Chojnacki et al., 2017).

Gene and repeat density plots for chromosomes were generated using karyoploteR (Gel and Serra, 2017). Secreted proteins were predicted using the neural network predictor of SignalP 3.0 (Bendtsen et al., 2004) and retained if they lacked a transmembrane domain outside the first 60 amino acids using TMHMM (Krogh et al., 2001). RepeatMasker 4.0.6 with the species fungi was used to softmask repeats. Repeats longer than 200 bp were used in the chromosome plotting.

Orthology analysis

Gene annotations with multiple isoforms were reduced to a representative isoform by selecting the longest CDS using a custom perl script. Orthologous proteins were identified with Orthofinder (Emms and Kelly, 2015) using default parameters. Multiple pairwise orthology analyses were run based on within-isolate and cross-isolate comparisons of similar haplotypes (i.e. A versus A or B versus C). Additional comparisons

were made between *Pgt*21-0 A, *Pgt*21-0 B, and Ug99 C haplotypes, as well as between Ug99 A, Ug99 C, and *Pgt*21-0 B haplotypes.

Phylogenetic analysis of rust isolates

For whole-genome SNP calling and phylogenetic analysis we used Illumina DNA sequence data (Supplementary Table 17) from the five Ug99 lineage isolates described here, seven Australian isolates we described previously (Upadhyaya et al., 2015; Chen et al., 2017) as well as 31 global isolates (Lewis et al., 2018) downloaded from the European Nucleotide Archive (ENA; PRJEB22223). Read quality was checked using FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and reads were trimmed with Trimmomatic v.0.33 (Bolger et al., 2014) using default settings and reads < 80 bp were discarded. Trimmed reads were aligned to the Ug99 or *Pgt*21-0 genome assemblies using BWA-MEM v.0.7.17 (Li and Durbin, 2009) and technical replicates were merged using SAMtools 1.6 (Li et al., 2009) and the PICARD toolkit (http://broadinstitute.github.io/picard/). Read lengths and coverage were verified by the functions bamtobed and coverage in BEDtools (Quinlan and Hall, 2010) and flagstat in SAMtools. FreeBayes v. 1.1.0 (Garrison and Marth, 2012) was used to call biallelic SNP variants across the 43 samples simultaneously. VCF files were filtered using vcffilter in vcflib (v1.0.0-rc1) with the parameters: QUAL > 20 & QUAL / AO > 10 & SAF > 0 & SAR > 0 & RPR > 1 & RPL > 1 & AC > 0. To verify that each sample consisted of a single genotype free of contamination, read allele frequencies at heterozygous positions were examined using the vcfR package (Knaus and Grünwald, 2017). VCF files were converted

to multiple sequence alignment in PHYLIP format using the vcf2phylip script (*https://zenodo.org/record/1257058#.XNnE845Kh3g*) and R-package ips/phyloch wrappings (*http://www.christophheibl.de/Rpackages.html*). Phylogenetic trees were constructed using the maximum likelihood criterion in RAxML v. 8.2.1.pthread (Stamatakis, 2006), assuming unlinked loci and using 500 bootstrap replicates with a general time reversible model. Convergence and posterior bootstopping (bootstrapping and convergence criterion) were confirmed with the *-I* parameter in RAxML and also with R-packages ape (Paradis et al., 2004), ips/phyloch, and phangorn (Schliep, 2010). Dendrograms were drawn using ggplot*2* (Wickham, 2016) and ggbio (Yin et al., 2012a) R-packages.

SNPs called against the A, B, or C haplotypes were separated from the total SNP sets using the function intersect –header in BEDtools. Homozygous and heterozygous SNPs were extracted by vcffilter -f "TYPE = snp" and -f "AC > 0 & AC = 2" and -f "AC > 0 & AC = 1", respectively and their frequency was counted using vcfkeepsamples and vcffixup. SNP statistics were calculated using vcfstats of VCFlib (v1.0.0-rc1).

Results

Haplotype-phased genome assembly

A single pustule derived from the original Ug99 isolate of *Pgt* (Pretorius et al., 2000) was purified and its pathotype was confirmed using the standard wheat differential set (Supplementary Table 1). We generated polished long-read genome assemblies for both

Ug99 and the previously described Australian stem rust isolate *Pgt*21-0 (Upadhyaya et al., 2015) using single-molecule real-time (SMRT) and Illumina sequence data (Supplementary Tables 2 and 3). Genome assembly of *Pgt*21-0 resulted in 410 contigs with a total size of 177 Mbp and a contig N_{50} of 1.26 Mbp (Supplementary Table 4). Similarly, the size of the genome assembly of Ug99 was 176 Mbp represented in 514 contigs with contig N_{50} of 0.97 Mbp. Thus, both assemblies were twice the size of a collapsed haploid assembly previously generated for a North American *Pgt* isolate (88 Mbp) (Duplessis et al., 2011), suggesting that in each assembly the sequences of the two haploid karyons were fully represented independently. Consistent with this, both genome assemblies contained over 96% of conserved fungal BUSCO genes, with the majority present in two copies (Supplementary Table 4). Furthermore, the *Pgt*21-0 assembly contained 69 telomeres, out of a total of 72 expected for a dikaryotic genome given the known haploid chromosome number of eighteen (Boehm et al., 1992). To identify sequences representing alternate haplotypes within each assembly we developed a gene synteny approach to assign contigs to groups representing paired homologous haplotypes (Figure 2.1). Using this approach, homologous pairs of sequences from each haplotype were assigned to 44 bins in *Pgt*21-0 and 62 bins in Ug99, which represented over 94% of each assembly (Supplementary Tables 4 and 5). Three of the eighteen chromosomes in *Pgt*21-0, and two in Ug99 seemed to be fully assembled as these bins contained telomere sequences at each end.

The *AvrSr50* and *AvrSr35* genes encode dominant avirulence factors recognized by wheat resistance genes (Chen et al., 2017; Salcedo et al., 2017). These two genes are located in close proximity (<15 kbp) to each other in the genome assemblies and both

haplotypes of this locus were assembled as alternate contigs in *Pgt*21-0 and Ug99 (Figure 2.2a). Both isolates were heterozygous for *AvrSr50* with one allele containing a ~26 kbp-insertion. *Pgt*21-0 was also heterozygous for *AvrSr35*, with one allele containing a 400 bp-MITE insertion previously described (Salcedo et al., 2017). PCR amplification had previously identified only a single *AvrSr35* allele, suggesting that it was homozygous in Ug99 (Salcedo et al., 2017). However, the Ug99 genome assembly contained a second allele of this gene, which contained a 57 kbp-insertion that would have prevented its PCR amplification in the Salcedo et al. study (Salcedo et al., 2017). The presence of this insertion in Ug99 was supported by DNA read (PacBio and Illumina) alignments across this genomic region and confirmed by DNA amplification and amplicon sequencing of flanking border regions (Supplementary Figure 1). Thus, Ug99 is also heterozygous for avirulence on *Sr35*, and may therefore mutate to virulence on this wheat resistance gene more readily than if it were homozygous. This is an important finding as it will inform *Sr35* deployment strategies against Ug99. Strikingly, the *AvrSr35-virSr50* haplotype of this locus is very similar in structure in Ug99 and *Pgt*21-0 (Figure 2.2a) and shares >99% sequence identity, while the two alternate haplotypes are quite different. We therefore compared the larger genomic regions containing these loci in each isolate: namely bin 06 in *Pgt*21-0, which is ~3.5 Mbp and includes telomeres at both ends, and bins 15 (1.8 Mbp) and 23 (1.2 Mbp) in Ug99 (Supplementary Figure 2a). One haplotype (designated A) was >99.7% identical in Ug99 and *Pgt*21-0, while the other two haplotypes (B and C) were highly divergent from each other and from haplotype A (Figure 2.2b, Supplementary Figure 2b and Supplementary Table 6). Only 71-78% of the sequences from each haplotype aligned, with an average identity of ~95% across the aligned regions, yielding total

identities of only 68-76% between these three distinct haplotypes. The high similarity between the A haplotypes of this chromosome suggested that Ug99 and *Pgt*21-0 may share large portions of their genomes, potentially up to an entire haploid genome copy.

Whole-genome haplotype assignment and comparison

We used a read subtraction and mapping approach (Figure 2.3a, b) to identify genome regions in the Ug99 and *Pgt*21-0 assemblies that showed high similarity and may be derived from a shared haplotype. Illumina reads from each isolate were mapped to the genome reference of the other isolate. Reads that failed to map were retained, and this subtracted read set, in which sequences common to both isolates had been removed, were re-mapped to their original genome reference. The subtracted read coverage depth was compared to the coverage depth obtained when using all reads (Supplementary Figure 3a, b). Contigs representing sequences shared by both isolates (designated as haplotype A) displayed a very low subtracted read coverage depth (Supplementary Table 7). In contrast, sequences unique to *Pgt*21-0 or Ug99 (designated as haplotypes B or C, respectively) retained a relatively high subtracted read coverage depth. Some contigs in each assembly appeared to be chimeric with distinct regions assigned to opposite haplotypes, and these contigs were divided into separate fragments (Supplementary Table 8) for subsequent haplotype comparisons. Approximately half of each genome assembly was assigned to either the A, B or C haplotypes (Figure 2.3c) and importantly one set of homologous sequences from each bin was assigned to each haplotype (Supplementary Table 7). The A, B and C haplotype sets contained 95-96% of conserved fungal BUSCO genes (Figure 2.3d),

indicating that each represents a full haploid genome equivalent. Consistent with this, the haplotypes were highly contiguous (Figure 2.4). Overall sequence identity between the A haplotypes of *Pgt*21-0 and Ug99 was over 99.5%, with structural variation (large insertions/deletions) representing only 0.46% of the haplotype sizes, with just 0.08% sequence divergence in aligned regions (Figure 2.4a, Table 2.1, and Supplementary Table 9). In contrast, only 91-93% of the A, B or C haplotypes could be aligned to each other, with an average sequence identity of ~96% across the aligned regions, giving total identities of 87% to 91% (Figure 2.4b to d, Table 2.1 and Supplementary Table 9). Structural variation between these haplotypes accounted for 6.7% to 8.6% of the haploid genome sizes. There were only ~9,000 SNPs (0.1/kbp) between the two A haplotypes, versus 876,000 to 1.4 million SNPs (11-18/kbp) between the A, B and C pairs, which is consistent with estimates of heterozygosity levels in *Pgt*21-0 (haplotypes A and B) based on variant detection from Illumina read mapping (Upadhyaya et al., 2015). The high similarity between the A haplotypes, and divergence between A, B and C haplotypes was also supported by Illumina read coverage and SNP calling analysis (Supplementary Figure 3c-e) showing that Ug99 and *Pgt*21-0 share one nearly identical haploid genome copy.

Assessment of inter-nuclear recombination

We tested two hypotheses that could explain the shared haplotype between Ug99 and *Pgt*21-0: 1) Ug99 arose by a somatic hybridisation event in which an isolate of the race 21 lineage donated an intact nucleus of the A haplotype (Figure 2.5a); and 2) Ug99 arose by a sexual cross in which one haploid pycnial parent was derived from a race 21 lineage

isolate after meiosis (Figure 2.5b). Under both scenarios, the A haplotype of Ug99 represents one entire haploid nucleus that was derived from the race 21 lineage isolate. In the nuclear exchange scenario, the *Pgt*21-0 A haplotype represents a single nucleus donated intact to generate Ug99. However, under the sexual cross model, this *Pgt*21-0 haplotype would include segments of both nuclear genomes that were combined by crossing over and chromosome reassortment after karyogamy and meiosis. Although recombination frequency has not been measured in *Pgt*, an average of 115 recombination events per haploid genome was detected during meiosis in the related flax rust fungus (*Melampsora lini*), which also has 18 chromosomes (Anderson et al., 2016). Because the *Pgt*21-0 and Ug99 genome assemblies represent the phased dikaryotic state of each isolate, all correctly phased contigs in Ug99 should be either A or C haplotype, while those in *Pgt*21-0 would include mixed haplotype contigs only if the sexual cross hypothesis is correct. In fact, just 19 contigs in the Ug99 assembly contained adjacent regions of either the A or C haplotype. These appeared to result from haplotype phase swap artefacts during genome assembly, since all of the junctions occurred between phase blocks (i.e. opposite gaps between the corresponding alternate contigs). Furthermore, examination of Illumina read mapping to these regions revealed that these sites contained either collapsed haplotype, non-unique sequences or discontinuities in read coverage (Supplementary Figure 4), indicative of assembly errors disrupting phase information across the junction. Likewise, 31 contigs of mixed haplotype in the *Pgt*21-0 assembly all contained likely phase swap artefacts (Supplementary Figure 4). To experimentally distinguish between phase-swap assembly artefacts and meiotic recombination events, we used Hi-C chromatin cross-linking proximity analysis (Belton et al., 2012) to assess physical linkage between contigs in the

*Pgt*21-0 assembly. About 90% of all read pair connections were between contigs of the same haplotype. For each of the chimeric contigs, the separated A and B fragments showed significantly more connections to contigs of the same haplotype than to contigs of the other haplotype, including other fragments of the original chimeric contig (Supplementary Table 10). These observed physical linkages confirmed that all of the mixed haplotype contigs in *Pgt*21-0 resulted from phase swap errors during the genome assembly process, and do not correspond to sites of genetic recombination that could have occurred during meiosis under the sexual origin hypothesis. Furthermore, only four structural variants larger than 10 kbp were detected between the A haplotypes of Ug99 and *Pgt*21-0. Three of these were deletions in Ug99, while the fourth was a tandem repeat duplication, indicating that there was no novel genetic content in Ug99 haplotype A compared to *Pgt*21-0.

Chromosome assembly and assessment of reassortment

Combining Hi-C scaffolding data with the bin and haplotype assignment information for the *Pgt*21-0 assembly allowed us to construct 18 chromosome pseudomolecules for each of the A and B haplotypes (Figure 2.6a, Supplementary Table 11 and 12). These covered a total of 170 Mbp and ranged from 2.8 to 7.3 Mbp in size, consistent with relative chromosome sizes from karyotype analysis (Boehm et al., 1992). The A and B chromosomes were collinear except for two translocation events (Figure 2.6b). In each case these were supported by contigs that spanned the translocation breakpoints. Re-scaffolding the separated fragments of these contigs using Hi-C data supported the original contig assembly, indicating that these are true translocation events within the A or

B genomes. The haplotype A chromosomes showed high collinearity with the Ug99 A haplotype contigs (Figure 2.6c).

Approximately 65% of the total Hi-C read pairs represented links between physically contiguous sequences on the same chromosome, while the remaining pairs connected sites distributed across the genome. Because Hi-C DNA crosslinking is performed in intact cells, these non-scaffolding linkages should preferentially form between chromosomes that are located in the same nucleus. Indeed, all chromosomes of the A haplotype showed a much higher proportion of Hi-C read pair links to other chromosomes of the A haplotype (~85%) than to chromosomes of the B haplotype (~15%) (Figure 2.6d), indicating that they are all located in the same nucleus. Thus, the A haplotype of Ug99 derives from a single nucleus of *Pgt*21-0. Similarly, 17 of the B haplotype chromosomes in *Pgt*21-0 showed stronger linkage to other B chromosomes (~90%) than to A chromosomes (~10%) (Figure 2.6e). However, chromosome 11B showed the inverse, suggesting that both homologs of this chromosome are in the same nucleus. This implies that a single chromosome exchange event occurred during asexual propagation of the *Pgt*21-0 isolate, after its divergence from the race 21 lineage branch giving rise to Ug99.

Overall the whole genome comparison data demonstrate that Ug99 shares one full haploid nuclear genome with the *Pgt*21-0 isolate with no recombination events within chromosomes and no reassortment of chromosomes from different nuclei. These facts are inconsistent with a sexual origin, and strongly support that the Ug99 lineage arose by a

somatic hybridisation event involving one parent derived from the African race 21 lineage and another parent of unknown origin exchanging whole nuclei (Figure 2.7).

Comparison of gene content between haplotypes

Annotation of the *Pgt*21-0 and Ug99 genome assemblies predicted similar gene numbers (~18,000) in each haplotype (Supplementary Table 13). Gene orthology analysis indicated that 65-70% of genes in each of the A, B and C haplotypes were shared and represent a core *Pgt* gene set, while the remainder were present in only one or two haplotypes (Supplementary Table 14). Mapping of orthologous gene pairs supported the overall synteny of the *Pgt*21-0 A and B chromosome assemblies and confirmed the translocations observed between chromosomes 3 and 5 and between 8 and 16 (Figure 2.8a). Genes encoding secreted and non-secreted proteins showed a similar distribution across the chromosomes, while repeat sequences displayed an inverse distribution to genes (Figure 2.8b Supplementary Figure 5). The location of secreted protein genes in gene-rich rather than repeat-rich regions is consistent with the absence of two-speed genome architecture in the related rust fungal species *P. coronata* and *P. striiformis* (Miller et al., 2018; Schwessinger et al., 2018).

Both Ug99 and *Pgt*21-0 are heterozygous at the predicted *a* and *b* mating type loci (Supplementary Figure 6), despite Ug99 being derived by a non-sexual mechanism. This is consistent with an expectation that formation and maintenance of a stable dikaryon requires two distinct compatible mating types (Park and Wellings, 2012). We observed

multiple alleles at the *b* locus on chromosome 9, which encodes the divergently transcribed transcription factors bE and bW, with variants sharing 70-80% amino acid identity. Cuomo and colleagues (Cuomo et al., 2017) previously described two alleles, *b1* and *b2*, for this locus in *Pgt*. *Pgt*21-0 contained the *b1* allele and a novel *b3* allele, while Ug99 contained *b3* and another novel *b4* allele. Both isolates were heterozygous for the + and – alleles of the locus on chromosome 4, which encodes a pheromone (mfa) and pheromone receptor (STE3) pair. However, one of the receptor alleles in Ug99 contained a single nucleotide deletion that resulted in truncation of the last 48 amino acids of the protein. Thus, the mating type system for *Pgt* appears to consist of two independent loci, one di-allelic and one multi-allelic.

Phylogenetic analysis of global *Pgt* isolates

We used the haplotype-phased genome references for *Pgt*21-0 and Ug99 to determine genetic relationships within a set of global *Pgt* isolates using publicly available sequence data (Upadhyaya et al., 2015; Chen et al., 2017; Lewis et al., 2018). Maximum likelihood trees based on whole genome SNPs (Figure 2.9a and Supplementary Figure 7a) showed a very similar overall topology to that reported previously for most of these isolates (Lewis et al., 2018). The five isolates of the Ug99 lineage, and the thirteen South African and Australian isolates each formed a separate tight clade, consistent with their proposed clonal nature (Park, 2007; Pretorius et al., 2007; Visser et al., 2019). However, tree building using filtered SNPs from just the A haplotype resulted in the formation of a single clade containing the Ug99, South African and Australian isolates, which indicates the clonal

derivation of this nucleus among these isolates (Figure 9b, Supplementary Figure 7b, c). The Ug99 group forms a subclade within the race 21 group consistent with a derived origin. In contrast, these groups remained in two distant clades in phylogenies inferred using filtered SNPs from the B genome. Surprisingly, in this case two isolates from the Czech Republic and three isolates from Pakistan were now located in a single clade with the South African and Australian isolates (Figure 2.9c). This suggests that these isolates contain a haplotype closely related to the B genome of the race 21 lineage and may also have arisen by somatic hybridisation and nuclear exchange. A phylogeny based on the C genome SNPs grouped isolate IR-01 from Iran with the Ug99 lineage (Figure 2.9d), suggesting that these isolates share the C haplotype. IR-01 could represent a member of the parental lineage that donated the C nucleus to Ug99, or alternatively may have acquired the C nucleus from Ug99. Notably, this was the only isolate that shared the *AvrSr35* 57kbp insertion allele identified in Ug99 (Supplementary Figure 8a). The relationships between these putative hybrid isolates were also supported by the patterns of homozygous and heterozygous SNPs detected in each haplotype (Supplementary Figure 8b-e). The incongruities between phylogenies generated based on different haplotypes highlight the difficulty of inferring relationships between isolates based on whole genome SNP data without haplotype resolution. Overall, these observations suggest that somatic hybridisation and nuclear exchange may be a common mechanism generating genetic diversity in global populations of *Pgt*.

Discussion

Although sexual reproduction of *Pgt* can generate individuals with novel genetic combinations, the completion of the sexual cycle requires infection of an alternate host, *Berberis* spp. (barberry) (Figueroa et al., 2016). In parts of the world where barberry is scarce or absent, either due to eradication programs or its natural distribution, *Pgt* is restricted to asexual propagation with new diversity often arising by mutation or migration (Park, 2007; Pretorius et al., 2007). Somatic hybridisation provides an alternative explanation for the appearance of new races not derived by stepwise mutation. Hybrids with high adaptive value in agroecosystems may establish new lineages of epidemiological significance, as shown by the emergence of the Ug99 lineage with its substantial impact on East African wheat production and threat to global food security (Singh et al., 2015). Results from experiments mostly conducted prior to the molecular biology era suggested the possibility of somatic genetic exchange between rust isolates co-infecting the same host, based on the detection of novel virulence phenotypes (Watson, 1957; Ellingboe, 1961; Flor, 1964; Bartos et al., 1969; Lei et al., 2017). In some cases only two non-parental phenotypic classes were observed, consistent with simple nuclear exchange (Flor, 1964; Bartos et al., 1969). The isolation of additional recombinant classes in other experiments was interpreted as recombination between nuclei, although this was also ascribed to isolate contamination (Flor, 1964; Bartos et al., 1969), which could not be ruled out without molecular data. Our data show that Ug99 arose by a somatic nuclear exchange event with no recombination. Phylogenetic analyses show that at least two different global lineages (from Pakistan and the Czech Republic) share a haplotype similar to the *Pgt*21-0 B genome, while another

(from Iran) shares a haplotype closely related to the Ug99 C genome. This suggests that multiple nuclear exchange events between strains have occurred in the global *Pgt* population and had significant impacts on genetic diversity. Although we did not observe any recombination between nuclei associated with the Ug99 hybridisation event, we did see evidence for translocation of one complete chromosome between nuclei in *Pgt*21-0. We also previously found that a *Pgt* mutant virulent on *Sr50* arose by exchange of an ~2.5Mbp region between two haplotypes (Chen et al., 2017). Thus, genetic exchange between haploid nuclei may occur as rare events during asexual propagation of a single lineage in *Pgt*. Whether extensive genetic exchange similar to ascomycete parasexuality (Fleißner and Serrano, 2016) can also occur between rust nuclei during hybridisation remains to be determined. This may require controlled infection experiments, as such recombinant hybrids would be difficult to distinguish from the products of sexual recombination in field derived-strains, especially given the potential for long-range spore dispersal.

Although there is now clear evidence of nuclear exchange between dikaryons in *Pgt*, nothing is known of how this process occurs or is regulated. It differs from parasexuality in ascomycetes (Fleißner and Serrano, 2016), as the dikaryotic state is maintained with no nuclear fusion or haploidisation resulting in chromosome reassortment. Wang and McCallum (Wang and McCallum, 2009) observed the formation of fusion bodies where germ tubes of different *P. triticina* isolates came into contact, with the potential for nuclear exchange at these junctions. It has been proposed that mating type loci contribute to determining the compatibility between isolates for the formation of hybrids

(Ellingboe, 1961; Flor, 1964), but this also remains to be confirmed experimentally. Our findings provide a new framework to take advantage of haplotype genome resolution to understand the role of somatic exchange in population diversity of rust fungi.

Extended dikaryotic developmental stages are common in many other fungi, especially basidiomycetes. Indeed, separation of karyogamy (fusion of haploid nuclei to form a diploid nucleus) from gamete fusion is a feature unique to the fungal kingdom (Stajich et al., 2009). However, it is unclear why fungi maintain an extended dikaryotic stage prior to formation of a diploid nucleus as a precursor to sexual reproduction (Kruzel and Hull, 2010). One possibility is that the ability to exchange haploid nuclei offers an advantage over the diploid state due to the enhanced genetic variation in long-lived asexual dikaryotes. There is also evidence for somatic exchange of genetic markers in dikaryotes of the mushroom *Schizophyllum commune*, which belongs to another Basidiomycete subphylum, Agaricomycotina (Clark and Anderson, 2004). Arbuscular mycorrhizae (AM) fungi are another ancient fungal lineage whose spores contain hundreds of nuclei and for which no sexual stages have been described, raising questions of how these lineages have survived (Corradi and Brachmann, 2017). Recently some dikaryotic-like AM isolates possessing two divergent classes of nuclei have been observed. Nuclear exchange between dikaryotes could be another driver of genetic variation in these fungi. Evidently, the members of the fungal kingdom display remarkable genetic plasticity and further investigation is required to reveal the mechanism, prevalence, and evolutionary importance of nuclear exchange in dikaryotic and multinucleate fungi.

Data availability

All sequence data, assemblies and gene annotation files generated in this study are available in NCBI under BioProject PRJNA516922. Assemblies and annotations are also available at the DOE-JGI Mycocosm Portal. Metadata for RNAseq libraries of *Pgt*21-0 and Illumina DNAseq libraries from all isolates are available in Supplementary Table 16 and 17, respectively. Data underlying Figure 2.3c and Supplementary Fig.3 a, b is provided in Supplementary Table 7. Data underlying Figure 2.6 is available in Supplementary Table 12. All other relevant data is available upon request from the corresponding authors.

Unless specified otherwise, all scripts and files are available at https://github.com/figueroalab/Pgt_genomes.

Table 2.1. Intra- and inter-isolate sequence comparison of entire haplotypes in Ug99 and *Puccinia graminis* f. sp. *tritici 21*-0.

	Sequence similarity		Structural variation		
				Total variant size	
Isolate comparison	**Bases aligned (%)**	**Sequence divergence (%)**	**Number of variants**	**Mbp**	**% of genome**
21-0 A vs Ug99 A	99.64	0.08	491	0.82	0.46
Ug99 A vs Ug99 C	91.52	4.08	2,571	13.69	7.88
21-0 A vs *Pgt*21-0 B	91.38	4.19	2,696	15.01	8.56
21-0 B vs Ug99 C	93.44	2.4	1,910	11.50	6.69

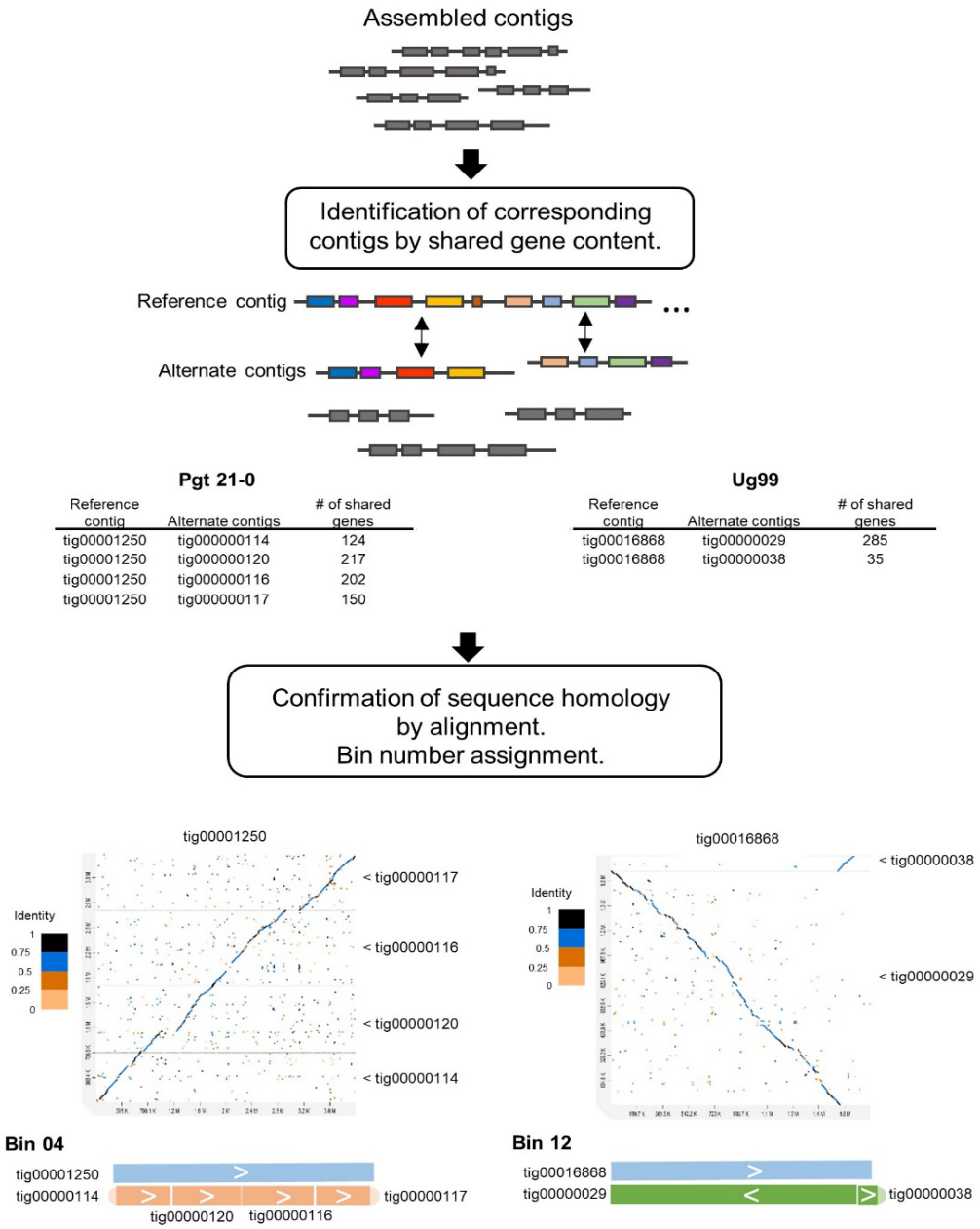


Figure 2.1. Strategy to identify homologous contigs in genome assemblies by gene synteny.

To detect shared content, Pgt gene models (Upadhyaya et al., 2015) (grey and coloured boxes) were aligned to the genome assemblies and the contig positions of the top two hits of each gene were recorded. Contigs containing at least five shared genes were considered as potential haplotype pairs. Sequence collinearity between contigs was assessed by alignment, and homologous matching contigs were assigned to bins. Examples shown are for Bin04 and Bin12 from Pgt21-0 and Ug99 respectively.

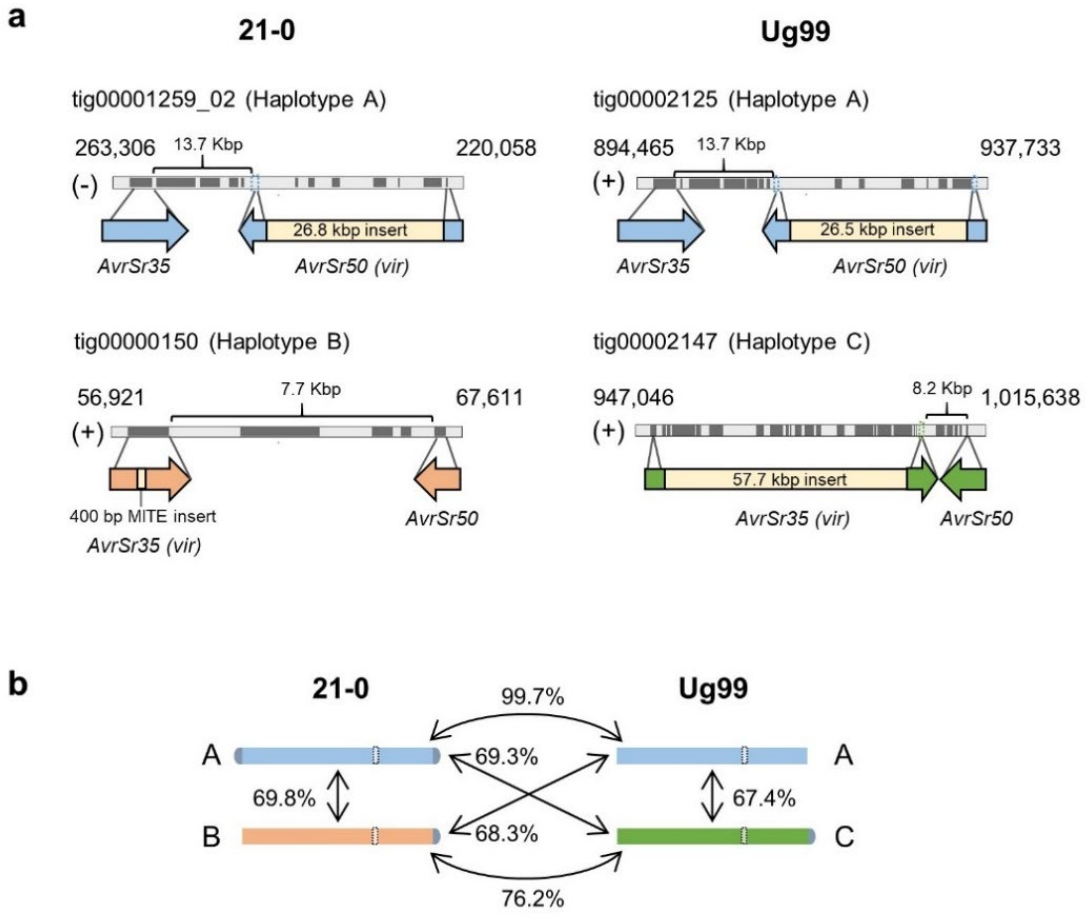


Figure 2.2. A common haplotype containing *AvrSr50* and *AvrSr35* is shared between *Pgt*21-0 and Ug99.

a) Diagram of genomic regions containing *AvrSr50* and *AvrSr35* alleles in *Pgt*21-0 and Ug99. Numbers above tracks correspond to contig coordinates and the sense of the DNA strand is indicated as + or -. Predicted gene models (including introns) are depicted as dark grey boxes and intergenic spaces are shown in light grey. Coloured arrows indicate location and direction of *AvrSr50* and *AvrSr35* genes, with size and position of insertions shown in yellow. Intergenic distances between *AvrSr50* and *AvrSr35* are indicated by brackets. **b)** Total sequence identity between contigs representing homologous chromosomes of different haplotypes (coloured bars) containing the *AvrSr50*/*AvrSr35* locus (dotted white boxes). Telomere sequences are represented in grey. Chromosome size = ~3.5 Mbp.

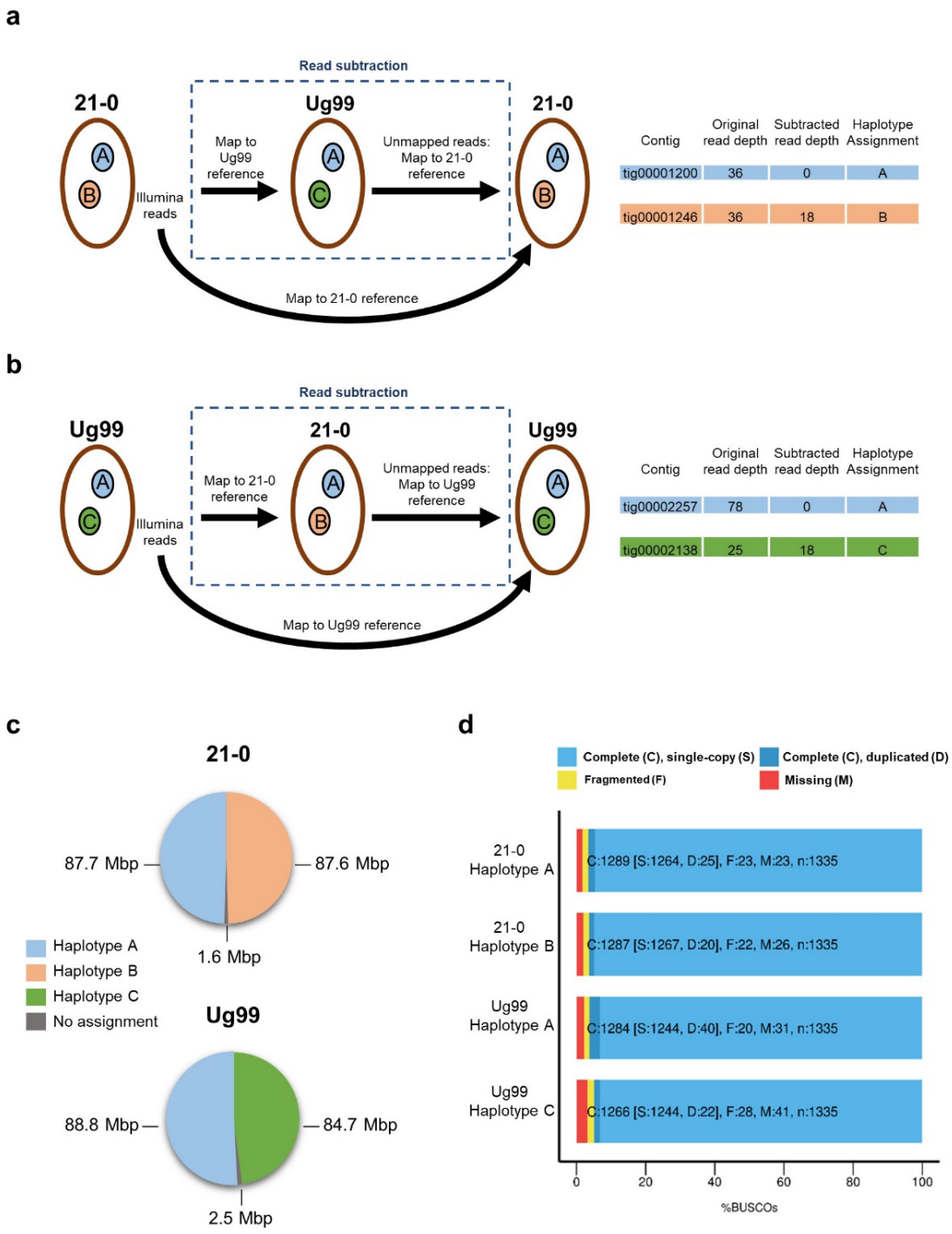


Figure 2.3. Haplotype assignment by read subtraction and mapping process.
a) Diagram of genomic regions containing *AvrSr50* and *AvrSr35* alleles in *Pgt*21-0 and Ug99. Numbers above tracks correspond to contig coordinates and the sense of the DNA strand is indicated as + or -. Predicted gene models (including introns) are depicted as dark

grey boxes and intergenic spaces are shown in light grey. Coloured arrows indicate location and direction of *AvrSr50* and *AvrSr35* genes, with size and position of insertions shown in yellow. Intergenic distances between *AvrSr50* and *AvrSr35* are indicated by brackets. **b)** Total sequence identity between contigs representing homologous chromosomes of different haplotypes (coloured bars) containing the *AvrSr50*/*AvrSr35* locus (dotted white boxes). Telomere sequences are represented in grey. Chromosome size = ~3.5 Mbp.

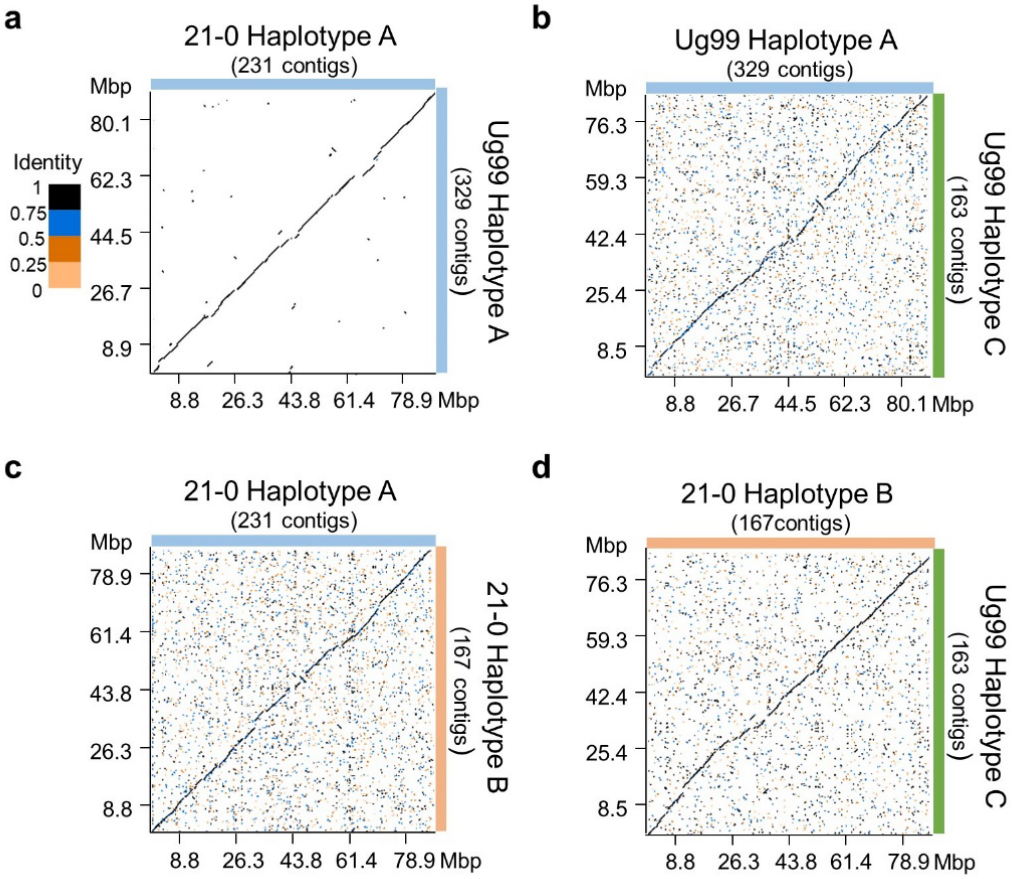


Figure 2.4. *Pgt*21-0 and Ug99 share one nearly identical haploid genome.
a) to **d)** Dot plots illustrating sequence alignment of complete haplotypes. X- and y-axes show cumulative size of the haplotype assemblies depicted by coloured bars to the right and top of the graphs. Colour key indicates sequence identity ratios for all dot plots.

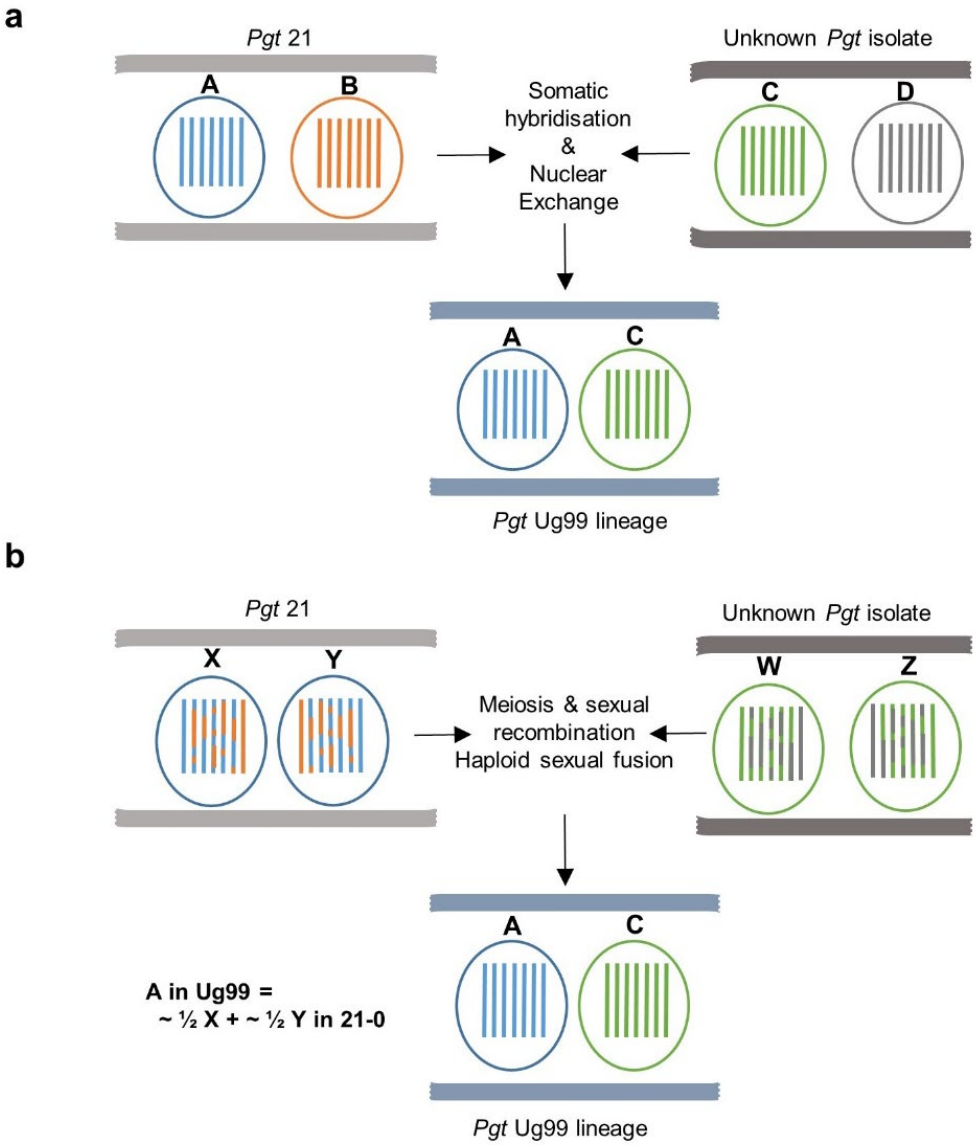


Figure 2.5. Models for the emergence of the founder isolate of the *Pgt* Ug99 lineage. **a)** A somatic hybridisation event and nuclear exchange occurred between an isolate of the *Pgt* 21 lineage and an unknown *Pgt* isolate. The combination of nuclei A and C yielded the parental isolate of the Ug99 lineage in Africa. Under this scenario, nucleus A of Ug99 is entirely derived from nucleus A in *Pgt*21-0. **b)** Alternatively, sexual reproduction and mating between these two parental isolates defined the origin of the Ug99 lineage. Under this scenario, meiotic recombination and chromosome reassortment would result in the *Pgt*21-0-derived A nucleus of Ug99 being composed of a mosaic of the two haploid nuclear genomes of *Pgt*21-0 (X and Y).

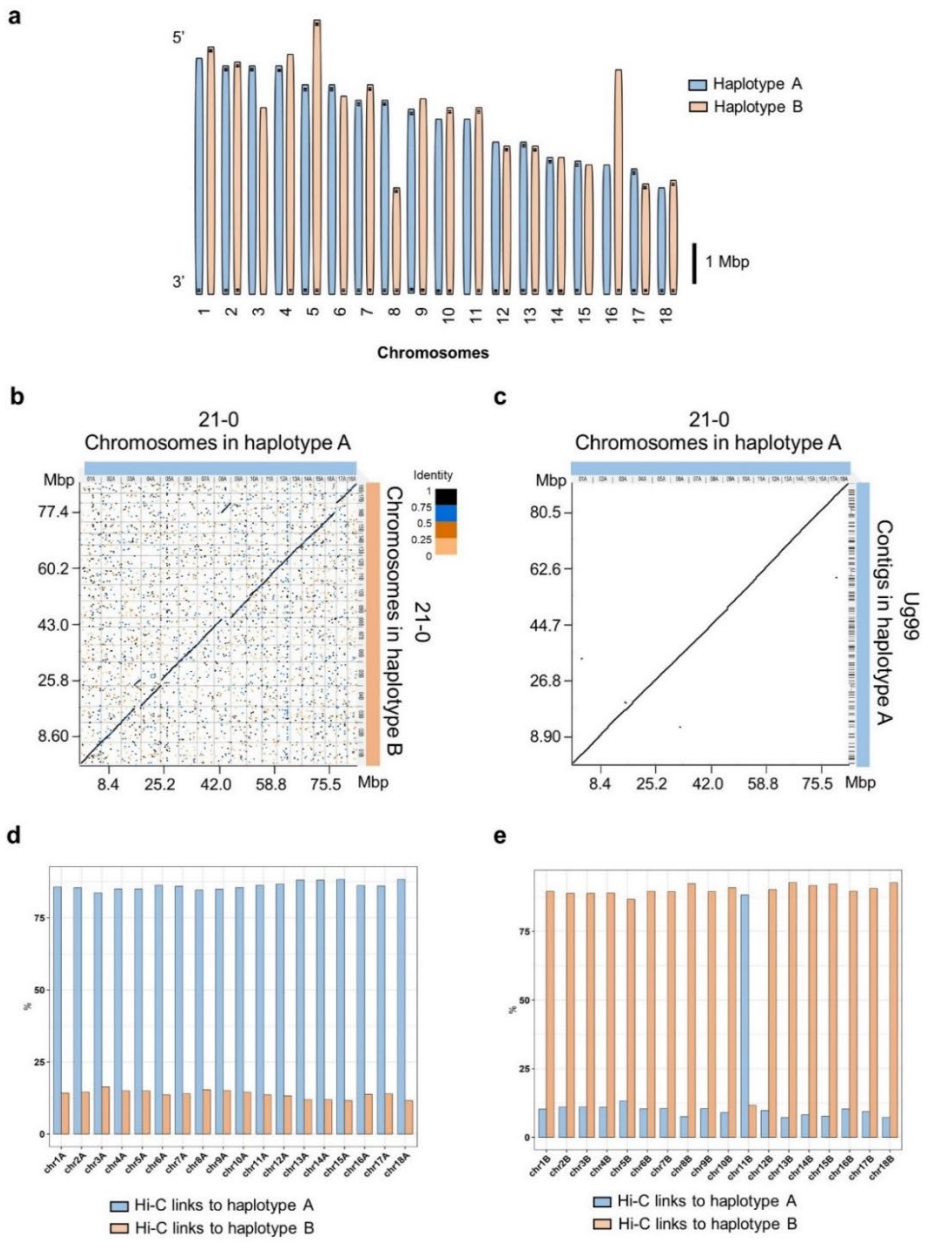


Figure 2.6. Chromosome sets of haplotype A and B in *Pgt*21-0.
a) Schematic representation of assembled chromosomes for *Pgt*21-0 of each haplotype (scale bar = 1 Mbp). Horizontal bars indicate telomeric repeat sequences. **b)** Dot plot of sequence alignment of *Pgt*21-0 chromosome pseudomolecules of haplotypes A and B. Two translocation events, one between chromosomes 3 and 5 and one between chromosomes 8 and 16, are evident. **c)** Dot plot of sequence alignment between chromosomes from haplotype A in *Pgt*21-0 and contigs from haplotype A in Ug99. **d)** Percentage of Hi-C read pairs linking each A haplotype chromosome to other A chromosomes A (blue) or to B haplotype chromosomes (orange). **e)** Percentage of Hi-C read pairs linking each B haplotype chromosome to either A (blue) or B (orange) chromosomes.

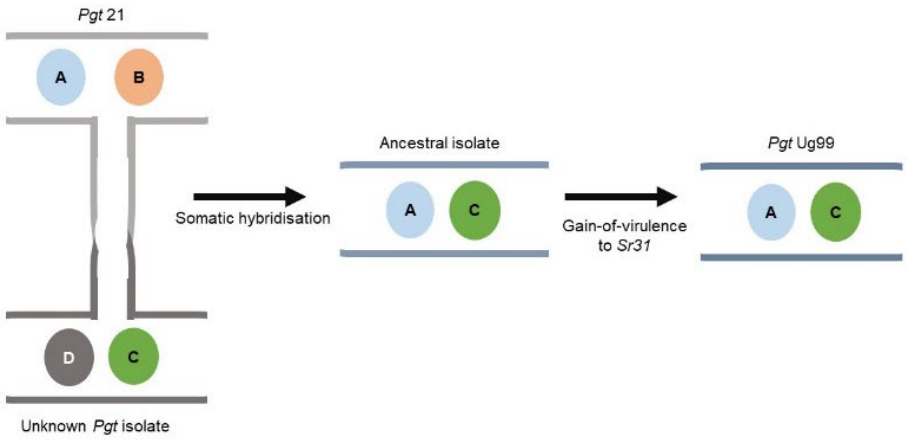


Figure 2.7. Model for Ug99 origin by somatic hybridisation and nuclear exchange between an isolate of the *Pgt* 21 lineage and an unknown *Pgt* isolate.
The ancestral isolate of the lineage acquired the A and C genomes and later gained virulence to wheat cultivars carrying the *Sr31* resistance gene.

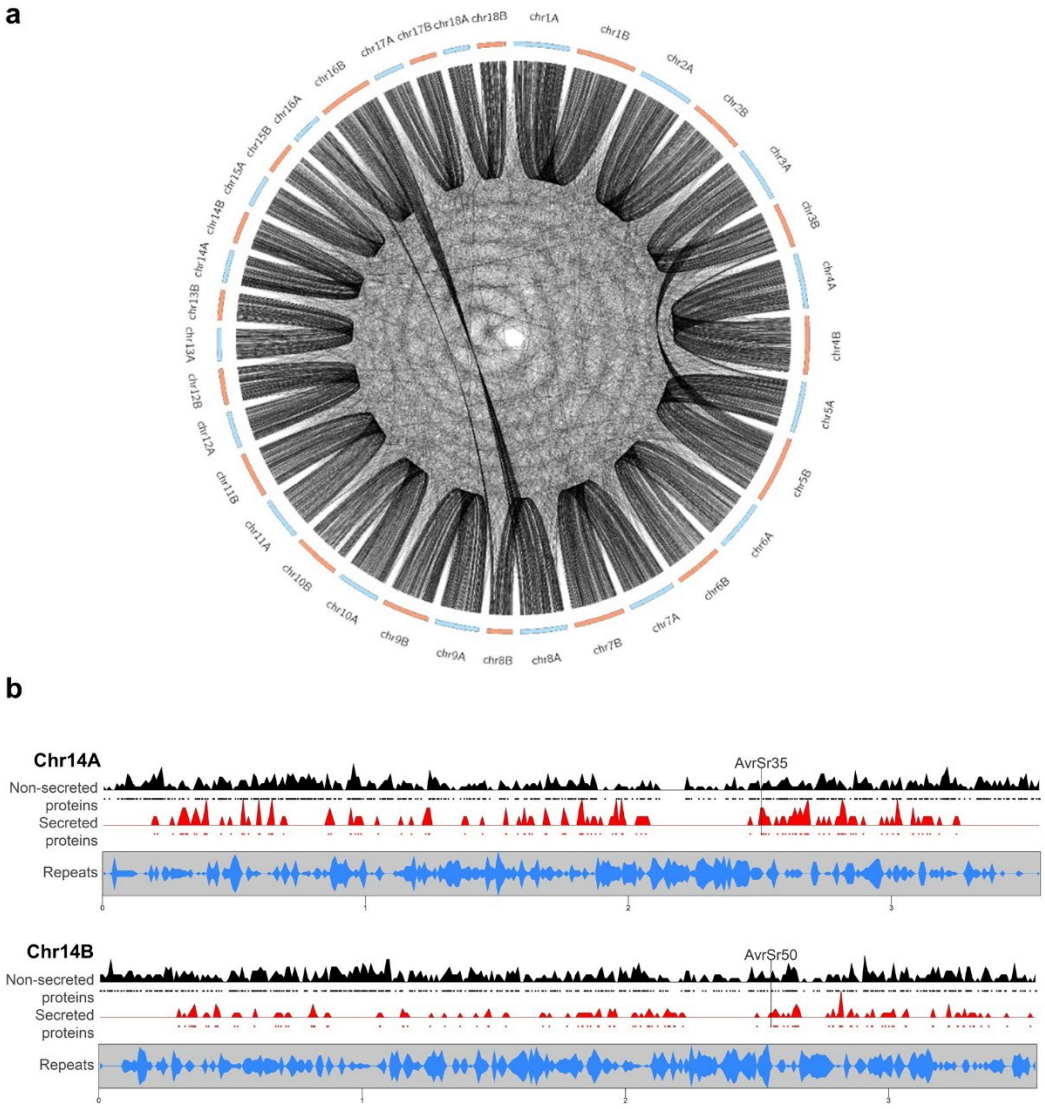


Figure 2.8. Gene content of *Pgt*21-0 chromosome pseudomolecules.
a) Circos plot showing location of orthologous gene pairs in the A and B chromosomes of *Pgt*21-0. **b)** Gene and repeat density plots for homologous chromosomes 14A and 14B. Density of genes encoding non-secreted (black) or secreted proteins (red) along the chromosomes are shown, with individual genes indicated by black or red dots. Bottom graph shows density of repeat elements (blue). Positions of *AvrSr50* and *AvrSr35* genes are indicated.

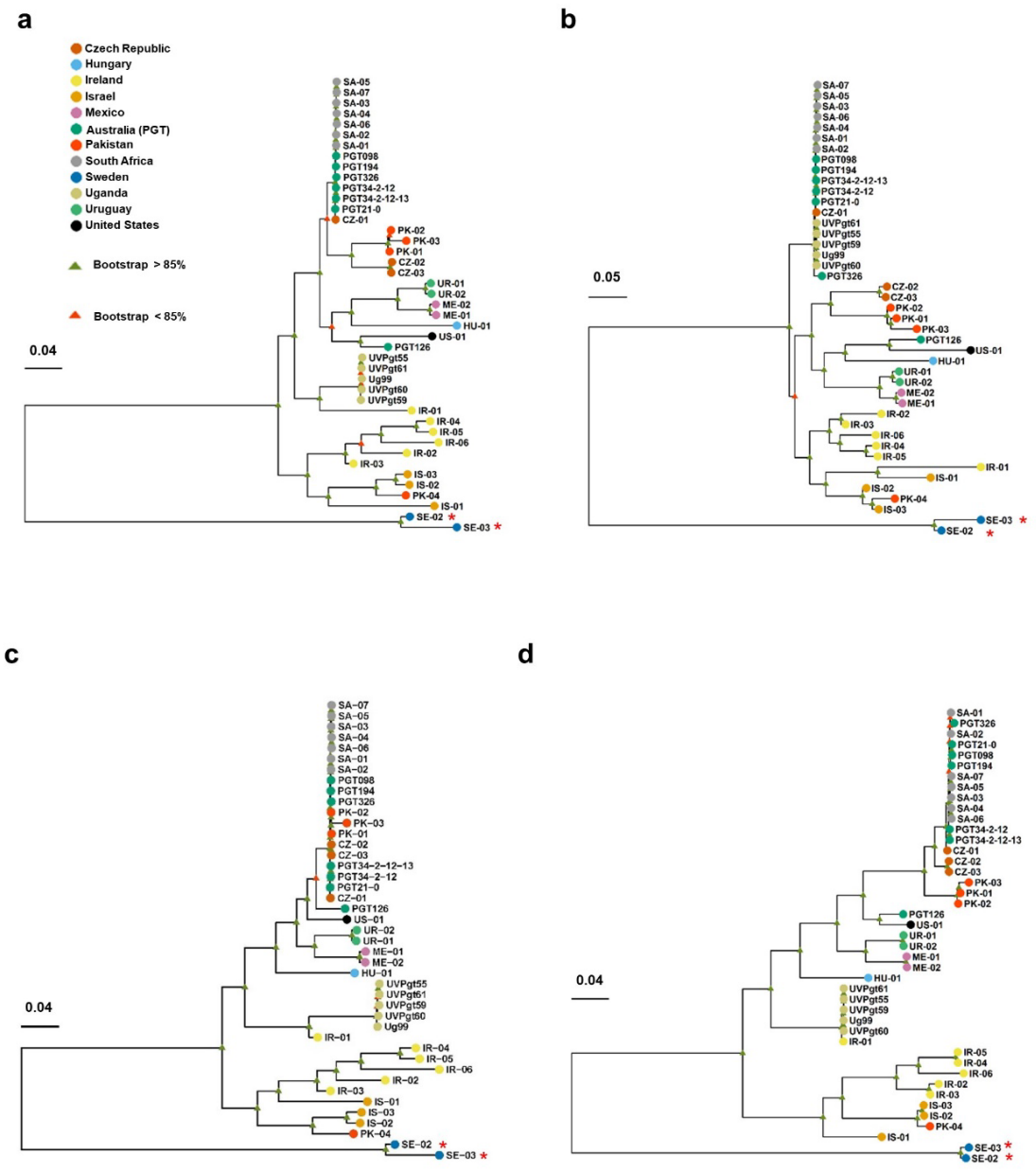


Figure 2.9. Somatic hybridisation in *Pgt* evolution.
a) Phylogenetic analysis of *Pgt* isolates from diverse countries of origin (colour key) using a RAxML model and biallelic SNPs called against the full dikaryotic genome of *Pgt*21-0. Scale bar indicates number of nucleotide substitutions per site. Red asterisks indicate *P. graminis* f. sp. *avenae* isolates used as an outgroup. **b)** Dendrogram inferred using biallelic SNPs detected against haplotype A of *Pgt*21-0. **c)** Dendrogram inferred using SNPs detected against haplotype B of *Pgt*21-0. **d)** Dendrogram inferred from SNPs detected in haplotype C of Ug99.

Chapter 3. Identification of genes mediating rust resistance or susceptibility in *Brachypodium distachyon*

Non-host resistance offers the possibility to provide durable resistance against a broad spectrum of pathogens. Here, I utilized a small grass and monocot model species *Brachypodium distachyon* as a system to study mechanisms of non-host resistance against *Puccinia graminis* f. sp. *tritici* (*Pgt*). Because of the phylogenetic relatedness of *B. distachyon* to wheat and barley, this species has potential applications for translational research to deliver durable and broad-spectrum stem rust resistance. In this study, a comparative RNA-Seq transcriptome analysis and a functional genomics approach assisted in the identification of candidate genes that may contribute to disease incompatibility/compatibility in *B. distachyon* challenged with the stem rust fungus. This study used a collection of T-DNA insertion mutants of *B. distachyon* to screen genetic factors mediating rust resistance and susceptibility. Here, two genes that may contribute to disease outcomes were identified for future validation.

Introduction

Incompatible pathogenic relationships in plant-microbe interactions are the result of a broad range of molecular and genetic events that are responsible for sensing biotic stressors and activating plant defense responses (Jones and Dangl, 2006). To withstand pathogen invasion, plants have evolved an elaborate multilayered microbial recognition

system involving two major strategies. One strategy involves PAMP-triggered immunity (PTI), which is prompted by the detection of conserved microbial elicitors known as pathogen-associated molecular patterns (PAMPS), or other plant-derived molecules resulting from cell damage, known as damage-associated molecular patterns (DAMPs) (Monaghan and Zipfel, 2012). In general, it is believed that non-adapted pathogens are unable to suppress PTI, a phenomenon that leads to non-host resistance (NHR); whereas adapted pathogens suppress PTI via the secretion of effectors that hijack and manipulate the host (Dodds and Rathjen, 2010). In the latter scenario, disease resistance can be achieved by a second and more specific strategy, effector-triggered immunity (ETI), which depends on the direct or indirect recognition of effectors by a cognate plant resistance (R) protein (Flor, 1971; Jones and Dangl, 2006; Dodds and Rathjen, 2010). On the other hand, the underlying mechanisms of NHR are poorly understood. In recent years, it has become evident that NHR involves mechanisms ranging from constitutive to inducible defenses, and contrary to conventional conceptions, NHR may involve both PTI- and ETI-associated immune responses (Schwessinger and Zipfel, 2008; Zipfel, 2008; Schulze-Lefert and Panstruga, 2011; Fan and Doerner, 2012; Omidvar et al., 2018). NHR has attracted increasing research interest due to its potential to provide broad-spectrum and durable resistance, and thus uncovering the molecular and genetic basis of NHR is a priority for its translational applications.

The causal agent of wheat stem rust, *Puccinia graminis* f. sp. *tritici* (*Pgt*), is responsible for significant yield losses of wheat and barley (Leonard and Szabo, 2005). A commonly used disease management method is to produce rust-resistant varieties of wheat

via introduction of *R* genes in elite cultivars (Ellis et al., 2014). Unfortunately, the efficacy of such approaches can be short-lived due to rapid shifts in virulence/pathogenicity in *Pgt* populations, as exemplified by the outbreaks of hypervirulent races such as Ug99 (TTKSK) as well as the often observed evolution of new virulence by single step mutations (Pretorius et al., 2000). Thus, the identification of plant species harboring natural resistance against stem rust is a continuous and crucial effort for crop protection programs. For example, researchers have been mined genetic resistance in rye (*Sr31*, *Sr50*) (Mettin et al., 1973; Zeller, 1973; Mago et al., 2015) and wild relatives of wheat such as *Aegilops tauschii* (*Sr33*, *Sr45*, *Sr46*) (Periyannan et al., 2013; Steuernagel et al., 2016; Arora et al., 2019) and *Triticum monococcum* (*Sr35*) (Saintenac et al., 2013). Alternatively, NHR using a close relative of wheat, *Brachypodium distachyon*, may provide a novel class of genetic materials to protect cereal crops against rust infections (Figueroa et al., 2015) as discussed in this Chapter.

B. distachyon (Poaceae), a model plant to study temperate grasses and cereals, displays natural diversity of immunity outcomes to the non-adapted pathogen *Pgt* (Ayliffe et al., 2013b; Figueroa et al., 2013, 2015). In fact, challenging *Brachypodium* accessions with *Pgt* led to development of various symptoms, from small flecks to expanding necrotic or chlorotic lesions and occasional tiny sporulating pustules, which is determined by genotype and environment (Ayliffe et al., 2013b; Figueroa et al., 2013, 2015). In addition to this, unlike rice which is a non-host to all rust species, *B. distachyon* is a compatible host to one rust pathogen, *Puccinia brachypodii* (Barbieri et al., 2011). *Brachypodium* is phylogenetically close to wheat, having diverged from the Triticeae 32 - 39 Mya and from

rice 40 - 53 Mya (International Brachypodium Initiative, 2010), and has been recently established as a model for studying non-host interactions of cereal rusts, specifically *Pgt*, *P. striiformis* f. sp. *tritici* (*Pst*) and *P. coronata* f. sp. *avenae* (*Pca*) (Ayliffe et al., 2013b; Figueroa et al., 2013, 2015; Dawson et al., 2015; Omidvar et al., 2018). *B. distachyon* is a diploid organism with five sets of chromosomes (2n = 10), a compact genome (272 Mbp) and short life cycle (~ 3 - 4 months) (International Brachypodium Initiative, 2010). These features make *Brachypodium* an attractive experimental model organism to study functionally important genes (Draper et al., 2001; Brkljacic et al., 2011). The development of genetic and genomic resources in *B. distachyon*, such as a collection of T-DNA insertion mutants generated by *Agrobacterium tumefaciens*-mediated transformation (Bragg et al., 2012), also makes this system suitable for functional genomics approaches to study genetic factors mediating responses to pathogens like *Pgt*.

The overall aim of this chapter was to identify and characterize putative genes underlying the resistance or susceptibility of *B. distachyon* to *Pgt* in light of its non-host status. A comparative RNA-Seq transcriptome analysis was performed in *B. distachyon* accession Bd1-1 to determine the early responses (12 and 18 hours post-inoculation (hpi)) to infection by *P. graminis* f. sp. *tritici* (*Pgt*), *P. graminis* f. sp. *phlei-pratensis* (*Pg-pp*), a pathogen of timothy grass, and *P. graminis* f. sp. *lolii* (*Pg-lo*), a pathogen of perennial ryegrass and tall fescue (Gordon et al., 2017, and unpublished data). After defining a list of potential genes of interest, I took a reverse genetics approach to validate the function of these genes in the interaction between *Pgt* and *Brachypodium*. As part of this work, I developed a mutant screening platform that involved phenotyping of *B. distachyon* T-DNA

mutant lines challenged with *Pgt* and genotyping to confirm the postulated mutation sites. Here, I present data suggesting that a WRKY transcription factor and the ortholog of the wheat gene *Lr67* (Moore et al., 2015) play a role in the compatibility outcomes between *Pgt* and *B. distachyon* accession line Bd21-3.

Materials and Methods

Plant materials and growth conditions

The *B. distachyon* inbred line Bd21-3 (Vogel and Hill, 2008) was obtained from the United States Department of Agriculture – Agricultural Research Service, Plant Science Unit (USDA-ARS) (St. Paul, MN, U.S.A.). T-DNA insertion mutants in the Bd21-3 genetic background (Supplemental Table 3.1) were acquired from the *Brachypodium* T-DNA Collection (Bragg et al., 2012) at the U.S. Department of Energy Joint Genome Institute (DOE-JGI) (Walnut Creek, CA, U.S.A.). To ensure uniform and synchronized germination, *B. distachyon* seeds were incubated at 4°C in the dark for 5 days on wet Whatman filter paper in an enclosed plastic bag or petri dish, and later shifted to room temperature (23-25°C) for 2 days before planting. For disease phenotyping, genotyping, and gene expression analysis of *B. distachyon* and wheat plants, seeds were planted at one seed per cone container (cone size 1 ¼" diameter X 8 ¼" depth) filled with Fafard® Germination Mix soil (Sun Gro Horticulture, Agawam, MA, U.S.A.). All the cone containers were placed on a 96-well cone rack and maintained in a greenhouse (16 hours light / 8 hours darkness; temperature range of 20-25°C and 30-50% humidity) before

inoculation. Seedlings were transplanted to Magnum pots (6”W X 6”L X 6.5” H) after 4 weeks if plants were kept for seed amplification. The wheat (*Triticum aestivum*) cultivar McNair, which serves as a susceptible host for *Pgt* in most cases, and the cultivated oat (*Avena sativa*) variety Marvelous used as a susceptible host to *P. coronata* f. sp. *avenae* (*Pca*) were both obtained from the USDA-ARS Cereal Disease Laboratory (USDA-ARS-CDL, St. Paul, MN, U.S.A.) and were maintained under the same greenhouse conditions mentioned above.

Fungal stocks and infection assays

Two rust fungal species were used in this study. Representative isolates of *Puccinia graminis* f. sp. *tritici* (*Pgt*): W1101 (race GFCDC, Figueroa et al., 2013) and 7a (also known as CDL75-36–700-3, race SCCL, Duplessis et al., 2011), and *Puccinia coronata* f. sp. *avenae* isolate (*Pca*) 12SD80 (race STTG, Miller et al., 2018) were first obtained from USDA-ARS-CDL and were maintained at -80°C as single-pustule cultures. The wheat cultivar McNair and oat cultivar Marvelous were used for spore amplification of *Pgt* and *Pca* isolates, respectively. Urediniospores were retrieved from -80°C storage and heat treated at 45°C for 15 min. Next, urediniospores from *Pgt* or *Pca* were suspended using mineral oil suspension (Soltrol 170, Chevron Phillips, Borger, TX, U.S.A..) at 14 mg/ml and 12 mg/ml, respectively. Seven-day-old wheat or oat seedlings with fully expanded primary leaves were spray-inoculated at the rate of ~ 50 µl/plant for infection assays using custom made spore inoculators (Tallgrass Solutions, Manhattan, KS) with a pump (Gast, Benton Harbor, MI). Inoculated plants were then transferred to a dark chamber under

continuous mist for 30 min and intermittent misting cycles for 16 h (4 min mist every 30 min). Subsequently, plants were exposed to light (150–250 µmol photons s^{-1} m^{-2}) for 4 h before transferring to a growth chamber under controlled conditions (18 h/6 h of light/dark, 24°C/18°C for day/night, 50% relative humidity). In parallel to the inoculation of plants, a 1% water agar plate was also sprayed with inoculum to record percentage of spore germination as indicative of pathogen viability. Germination rates between 80-100% were considered acceptable. For amplification of urediniospores, spores were collected at 9 and 14 days post inoculation (dpi), dehydrated in a desiccator containing desiccant-anhydrous indicating drierite (W. A. Hammond Drierite Co. Ltd, Xenia, OH, U.S.A.) and stored at -80°C.

B. distachyon plants were infected at the three-leaf stage (10-14 days after planting) following the same inoculation procedure. Each infection assay was accompanied by seven-day-old seedlings of wheat McNair or oat Marvelous as positive controls to monitor the disease symptom development and successful infection by *Pgt* or *Pca* accordingly. Plants were randomly placed on the cone rack before inoculation to prevent bias during inoculations.

Disease scoring of *Brachypodium distachyon* seedlings

Disease phenotyping was conducted in mutants at T_2 and T_3 generations following T_1 single-seed descent increases. Initial disease screening for alterations in disease resistance in the T-DNA insertion mutants was conducted by infection assays using the *Pgt*

isolate W1101 and with wild type (WT) Bd21-3 as reference control. Approximately 10 to 25 seedlings of each *B. distachyon* T-DNA insertion line and WT were inoculated for the initial disease screening. The number of seedlings depended on the amount of available seeds and seed germination rate as growth of some *B. distachyon* mutants was not as robust as that of WT seeds. Symptoms were evaluated at 14 days post inoculation (dpi). The number of macroscopic lesions including necrotic and chlorotic areas was recorded as a metric of disease severity, and the means of these values were calculated for downstream comparisons. T-DNA mutants that displayed changes in disease severity with respect to WT were shortlisted for further disease scoring in second and/or third trials.

To further examine susceptibility or resistance phenotypes for mutant lines of interest, symptom development in response to *Pgt* (isolate 7a and W1101) was recorded at 14 dpi for individuals of the mutant line JJ19596 (homozygous for T-DNA insertion, T2 generation) using macroscopic lesion count in one inoculation trial containing 8-10 plant replicates. Statistical analysis was performed using a Mann-Whitney test (Ruxton, 2006) to compare the phenotype changes between mutants and WT. Similarly, disease severity was recorded with responses to *Pgt* 7a and *Pca* 12SD80 for individuals of the mutant line JJ16783 (homozygous for T-DNA insertion, referred to as HM; T2 generation) by counting the macroscopic lesions at 14 dpi. This was performed in two independent trials with each containing 18-30 replicate seedlings for WT and the mutant line JJ16783. Here, the experiments had the unequal sample size for different genotypes in each trial: the total number of 26, 22, 21, 22 plants of infected WT, JJ16783-HM2, HM5, HM8 was assayed for Trial 1, and 30, 22, 19, 18 plants of infected WT, JJ16783-HM2, HM5, HM8 was

assayed for Trial 2. In addition, as the two inoculation trials were conducted at independent times, random trial effects needed to be considered. Therefore, to allow for the unbalanced design and random trial effects, statistical analysis was conducted using linear mixed-effects (lme) models in the lme4 package in R (Baayen et al., 2008; Bates et al., 2014) to test for the phenotypic differences between mutants and WT. Macroscopic lesions of infected leaves were captured using a stereomicroscope (Olympus model SZX16).

Microscopic analysis of fungal growth in *B. distachyon* seedlings

Fungal infection structure development and colonization of *Pgt* (isolate 7a and W1101) were assayed in WT Bd21-3 and the WRKY transcription factor *BdWRKY46* mutant (JJ19596, T2 plants) following previously described procedures (Dawson et al., 2015; Omidvar et al., 2018). Briefly, secondary leaves of infected *B. distachyon* and wheat plants were collected at 14 dpi and submerged in a 15 ml-centrifuge tube containing 6 ml of 1 M KOH for at least 12 h before neutralizing samples in 50 mM Tris (pH 7.5). Subsequently, leaf tissues were stained with wheat germ agglutinin conjugated to the fluorophore alexa Fluor® 488 (WGA-Alexa) (Invitrogen, Carlsbad, CA, U.S.A.) at 20 µg/ml. Samples were examined using a fluorescence microscope (Leica model DM LB; Leica, Houston, TX, U.S.A.) with a blue light of 450 - 490 nM excitation filter. The percentage of fungal colonization (pCOL) and pustule formation (pPUST) was quantified using the score scale as defined (Dawson et al., 2015; Omidvar et al., 2018). Score of 0, 0.5, 0.75 and 1 were assigned and recorded to each disjointed field of view (FOV) with hyphal growth covering less than 15%, 15-50%, 50-75% and over 75% of the visible area,

respectively, by scanning the whole leaf fragment. Final pCOL or pPSUT score was calculated by weighted average score of the total number of FOVs of each leaf. Data was collected from secondary leaves of 8-10 plant replicates for WT and the mutant lines in one inoculation trial. Statistical analysis was performed using Mann-Whitney test (Ruxton, 2006) to compare pCOL and pPUST results between WT and T-DNA insert homozygous mutant individuals.

DNA extraction and PCR-based genotyping for T-DNA insertion mutants

Genomic DNA extraction was performed as described in (Edwards et al., 1991) and modified as follows. Three leaves of each *Brachypodium* plant were pooled for total DNA extraction. For each sample, foliar tissue was ground in liquid nitrogen in a 1.5 ml eppendorf microcentrifuge tube using a pellet pestle (Sigma-Aldrich) before homogenizing with 500 µl of DNA extraction buffer (200 mM Tris-HCl (pH 8.0), 250 mM NaCl, 25 mM EDTA, 0.5% SDS). Next, each sample was mixed with 200 µl of chloroform:isoamyl alcohol (24:1) (Sigma-Aldrich) and centrifuged at 10,000 *g* for 10 min. The supernatant was transferred to a new 1.5 ml microcentrifuge tube and mixed with 300 µl of Isopropanol (Sigma-Aldrich). Samples were then incubated at room temperature (23-25°C) for 5 min and centrifuged at 10,000 *g* for 10 min. Each DNA pellet was rinsed with 500 µl of 75% ethanol, and resuspended in 100 µl of 1X TE buffer at 37°C for 1 h before storing at 4°C. DNA was quantified using a NanoDrop 2000 Spectrophotometer NanoDrop (Thermo Scientific, Waltham, MA, U.S.A.).

For the purpose of genotyping T-DNA insertion mutants, gene-specific primers for each candidate gene (Table 3.2) were designed using Primer3 (Untergasser et al., 2012) to amplify ~ 500 bp upstream and downstream of the postulated T-DNA insertion site (Supplemental Figure 3.1) according to the reference genome Bd21 assembly v3.1 in Phytozome (coordinates of T-DNA inserts and be found in this link https://jgi.doe.gov/our-science/science-programs/plant-genomics/brachypodium/brachypodium-t-dna-collection/). Gene-specific primers MF86, MF87 (Table 3.1) were used for amplification of *BdWRKY46* (Gene ID: *Bradi2g22440*) in JJ19596 (T1 generation). Gene-specific primers MF119, MF120 (Table 3.1) were used for amplification of *BdSTP13* (Gene ID: *Bradi1g69710*) in JJ16783 (T1 generation). Gene-specific primers were used in pairs or combined with a primer directed to the left border of T-DNA construct pJJ2LBA (T3 T-DNA LB) to target the wild-type region of the candidate gene and the T-DNA flanking chimeric region (which spans the insertion border between the gene and the T-DNA), respectively. PCR reactions were set up using a Premix Taq™ DNA Polymerase (TaKaRa Taq™ Version 2.0) (Takara Bio U.S.A., Inc., Mountain View, CA) following the manufacturer's protocol. The PCR was conducted in a C1000 Touch™ Thermal Cycler (Bio-Rad, Hercules, CA, U.S.A.) using the following program: initial denaturation at 98°C for 30 s, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 63°C for 30 s, extension at 72°C for 1 min 20 s, and a final extension at 72°C for 10 min. PCR products were later separated using 1% agarose gel electrophoresis in 1X Tris-Borate-EDTA (TBE) Buffer and the GeneRuler 1kb DNA Ladder Plus (Thermo Scientific) was used for sizing of amplicons. PCR amplicons were visualized with a SYBR Safe DNA gel stain (Invitrogen) in a Gel Doc 2000 Gel Imager (Bio-Rad) using the transilluminator setting.

PCR products were subsequently purified using an UltraClean® PCR clean-up kit (MO BIO Laboratories Inc, Carlsbad, CA, U.S.A.) for Sanger sequencing at the University of Minnesota Genomics Center (UMGC) (Table 3.1). Base calling was performed using Sequencher 4.10.1 DNA sequence analysis software (Gene Codes Corporation, Ann Arbor, MI, U.S.A.), and sequences were aligned using Clustal Omega (Chojnacki et al., 2017) to confirm T-DNA insertion sites and homozygosity or heterozygosity of mutations.

RNA isolation and cDNA synthesis

Tissues for RNA isolation were collected from four secondary leaves of WT Bd21-3 or mutant lines in MP Lysing Matrix E 2 ml tubes (MP Biomedicals, Solon, OH, U.S.A.) in liquid nitrogen. Wheat tissues for RNA isolation were collected from three primary leaves and pooled together in the Lysing Matrix E tubes. Next, total RNA was extracted using RNeasy Plant Mini Kit (Qiagen, Germantown, MD, U.S.A.) according to the manufacturer's instructions, and then treated with RNase-free DNase set (Qiagen) to remove DNA from samples. RNA was quantified using a NanoDrop 2000 Spectrophotometer NanoDrop (Thermo Scientific), and the integrity of RNA was assayed using Agilent 2100 Bioanalyzer at UMGC. Next, 1st stranded cDNA was synthesized from 500 ng RNA using a PrimeScriptTM 1st strand cDNA Synthesis Kit (Takara) and following the manufacturer's instructions in a C1000 Touch™ Thermal Cycler (Bio-Rad).

Time-course gene expression analysis of *BdSTP13*

WT Bd21-3 was infected with *Pgt* isolate 7a or *Pca* isolate 12SD80, and leaf tissues were collected for RNA extraction at 1, 2, 3, 4, 7, 9, 11, and 13 dpi, respectively. Four leaves were pooled together for RNA extraction as one biological replicate. Time-course gene expression of *BdSTP13* (*Bradi1g69710*) was quantified using reverse-transcription quantitative PCR (RT-qPCR) in a CFX96 Real-Time PCR system (Bio-Rad). To measure gene expression of *BdSTP13* (*Bradi1g69710*), RT-qPCR was carried out using a 1:4 dilution of 1st stranded cDNA template, primers MF125, MF126 (Table 3.1) and SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) following the manufacturer's protocol. The thermal cycling program included a denaturation step at 98°C for 30 seconds, 35 cycles of denaturation at 98°C for 10 seconds, annealing/extension at 60°C for 30 seconds. A no-template control was used as a negative control. Quantification PCR (qPCR) efficiency for MF125 and MF126 is 118% (Supplemental Figure 3.1). *BdSTP13* expression fold changes were calculated relative to mock-treated plants in a comparative C_T method ($\Delta\Delta C_T$ method) (Livak and Schmittgen, 2001; Omidvar et al., 2018), using the *B. distachyon GAPDH* (Gene ID: *Bradi3g14120*) as reference gene to normalize expression (amplified using primers BdGAPDH_F and BdGAPDH_R) (Table 3.1). In addition, wheat McNair plants were inoculated with *Pgt* 7a, and expression profile of *Lr67sus* (4D hexose transporter, Lr67_DD, NCBI accession: KR604816) was recorded in the same time-course as described for WT Bd21-3. RT-qPCR experiments were set up as described above using the primers Lr67_DD_F and Lr67_DD_R (Moore et al., 2015) (Table 3.1). *Lr67sus* expression fold changes were calculated relative to mock-treated plants using $\Delta\Delta C_T$

method, with the wheat *GAPDH* (EU022331) as reference gene which was amplified using primers MF142 and MF143 (Yin et al., 2009) (Table 3.1). For all these experiments, the mean and standard error of the mean (SEM) of the relative expression at each time point were calculated based on data from three biological replicates. Statistical analysis was performed using one-way ANOVA followed by multiple comparisons to compare the gene expression changes between infected plants and mock inoculation.

Detection of *BdSTP13* transcripts in Brachypodium T-DNA mutant

Leaf tissues from WT and mutant plants were collected at 4 dpi after *Pgt* 7a and *Pca* 12SD80 infection, followed by RNA extraction and cDNA synthesis as described above. Gene expression of *BdSTP13* (*Bradi1g69710*) was quantified using reverse-transcription PCR (RT-PCR) with cDNA as template and primers MF125, MF126 and MF129, MF130 (Table 3.1). The *B. distachyon GAPDH* (Gene ID: *Bradi3g14120*) was used as an internal reference gene and was amplified using primers BdGAPDH_F and BdGAPDH_R (Table 3.1). PCR amplifications were set up using Premix Taq™ DNA Polymerase following the manufacturer's instructions in C1000 Touch™ Thermal Cycler (Bio-Rad) in a program including an initial denaturation step at 98°C for 30 seconds, followed by 40 cycles of denaturation at 98°C for 10 seconds, annealing at 60°C for 30 seconds, extension at 72°C for 24 seconds, and a final extension at 72°C for 10 min. RT-PCR products were later separated using 2% agarose gel electrophoresis in 1X TBE and stained with a SYBR Safe DNA gel stain (Invitrogen) and visualized using a Gel Doc 2000

Gel Imager (Bio-Rad) using a transilluminator setting. A 50 bp-DNA ladder (New England BioLabs Inc. Ipswich, MA, U.S.A.) was used as marker to estimate amplicon sizes.

Phylogenetic tree construction of members of the sugar transporter proteins family

The orthologs of 14 sugar transporter proteins (STP) members defined in *Arabidopsis* were investigated in *B. distachyon* using the protein sequences of *B. distachyon* (Bd21 assembly v3.1) and *Arabidopsis* (*Arabidopsis thaliana* TAIR10) in the program OrthoFinder v1.1.8 using default parameters (Emms and Kelly, 2015). The identifiers of the STP13-like members in various plant species (*Arabidopsis thaliana*, *Brachypodium*, *Hordeum vulgare*, *Lycopersicon esculentum*, *Oryza sativa*, *Sorghum bicolor*, *Setaria italica, Vitis vinifera*, *Zea mays*, *Triticum urartu*, *Aegilops tauschii*) were obtained from Moore et al. (2015). A maximum-likelihood phylogeny analysis of all STP13-like members was performed using a web-based program called NGPhylogeny.fr (https://ngphylogeny.fr/) (Lemoine et al., 2019). The phylogenetic tree was built with protein sequences aligned using Multiple Alignment using Fast Fourier Transform (MAFFT), Block Mapping and Gathering with Entropy (BMGE) curation, Smart Model Selection (PhyML+SMS) tree inference, and Newick tree rendering, all following default settings.

Cloning of *BdSTP13* and vector construction

Total RNA was extracted from leaves of *Pgt* 7a-treated Bd21-3 (WT) at 4 dpi using RNeasy Plant Mini Kit (Qiagen). The first strand cDNA was reverse transcribed using PrimeScript™ 1st strand cDNA Synthesis Kit (Takara). The full-length cDNA sequence (~1.6 kbp) of *BdSTP13* (*Bradi1g69710*) was PCR amplified using Phusion High-Fidelity DNA polymerase (New England BioLabs Inc.) with the primers MF113 and MF114 (Table 3.1) to add *EcoR*I and *Sal*I restriction sites flanking the start codon and the stop codon of *BdSTP13* cDNA. The PCR cycle conditions included 98°C for 30 s, followed by 30 cycles of 98°C for 10 seconds, 67°C for 30 seconds, 72°C for 50 seconds; 72°C for 10 min in a C1000 Touch™ Thermal Cycler (Bio-Rad). PCR amplicons were separated by 1% agarose gel electrophoresis in 1X TBE, and the *BdSTP13* amplicon (~1.6 kbp) was purified using a QIAquick® Gel extraction kit (Qiagen). The *BdSTP13* full-length cDNA insert and the plasmid pBluescript SK (-) (pBS) (Short et al., 1988) was both double digested with restriction enzymes *EcoR*I and *Sal*I (Thermo Scientific) following the manufacturer's instructions. The digested products were separated in a 1% agarose gel in 1X TBE, purified using a QIAquick® Gel extraction kit (Qiagen) and subsequently used in the ligation reaction between *BdSTP13* cDNA insert and plasmid pBS at the ratio of 3:1 (DNA insert:plasmid) using a Rapid DNA ligation Kit (Thermo Scientific). The recombinant pBS plasmid was transformed into Turbo Competent *E. coli* cells (New England BioLabs Inc.) following the manufacturer's instructions. Blue-white selection of transformants was carried out after overnight incubation at 37°C on 100 mg/ml ampicillin Luria-Bertani (LB) agar plates previously treated with 100 ml of 40 mM IPTG (Sigma-Aldrich) and 100 ml of

2% X-gal (Fisher Scientific). White colonies including recombinant plasmids (pBS-*BdSTP13*) were purified using Plasmid Mini kit (Qiagen). Recombinants were verified by double digestion of the purified plasmids using restriction enzymes *EcoR*I and *Sal*I as described above. In addition, the clones of *BdSTP13* were confirmed by Sanger Sequencing at UMGC using primers pBS_SK_M13_F and pBS_SK_M13_R (Table 3.1).

Following the same procedure, the *BdSTP13* cDNA was cloned into the yeast expression vector pDR196 (Rentsch et al., 1995), which contains the *URA3* gene for uracil prototrophy (Milne et al., 2017), and into the pDR196T backbone (Milne et al., 2019), which contains the *TRP* gene for tryptophan selection, respectively. The ligation products were transformed into *E. coli*, and transformants were selected on 100 mg/ml ampicillin LB-agar plates after overnight incubation at 37°C. Single colonies were purified as previously described. Constructs were verified by Sanger sequencing at UMGC using the primers MF127, MF128, pDR196_fPMA5 and pDR196_rADHclose (Table 3.1).

Site-directed mutagenesis of BdSTP13 and construction of vectors for heterologous expression in yeast

*Dpn*I-mediated site-directed mutagenesis was conducted in the pBS-*BdSTP13* construct to introduce the amino acid mutations G144R and V387L separately using a method described by Milne et al (2019). This approach was based on a modified version of the QuikChange protocol (Agilent, Mulgrave, VIC, Australia) for site-directed mutagenesis with primers designed at the QuikChange Primer Design webpage

(http://www.genomics.agilent.com/primerDesignProgram.jsp) based on the protein and CDS sequences of *BdSTP13* (*Bradi1g69710*). In brief, mutant strands were synthesized by PCR using Phusion High-Fidelity DNA polymerase (New England BioLabs Inc.) following the manufacturer's protocol. The primers BdSTP13_G144R_F and BdSTP13_G144R_R were used to generate $BdSTP13^{G144R}$ cDNA clone while BdSTP13_V387L_F and BdSTP13_V387L_R were used to generate the $BdSTP13^{V387L}$ cDNA clone (Table 3.1). The thermal cycle conditions included 98°C for 30 seconds; followed by 20 cycles of 98°C for 10 seconds, 55°C for 1 min, 72°C for 1 min 50 seconds; 72°C for 3 min using a C1000 Touch™ Thermal Cycler (Bio-Rad). PCR amplicons were directly purified using the UltraClean® PCR clean-up kit (MO BIO Laboratories Inc.), followed by *Dpn*1 digestion (New England BioLabs Inc.) that included a reaction of 2 µl of *Dpn*1 buffer, 1 µl of *Dpn*1 enzyme and 17 µl of purified PCR products and an incubation period of 2 h at 37°C. Site-directed mutant cDNA clones were amplified by transformation the recombinant vectors (pBS-$BdSTP13^{G144R}$ and pBS-$BdSTP13^{V387L}$) into *E. coli* cells followed by blue-white selection and plasmid purification as described earlier. The double mutant $BdSTP13^{G144R+V387L}$ was obtained by performing a second round of the QuikChange protocol in pBS-$BdSTP13^{G144R}$ using primers BdSTP13_V387L_F and BdSTP13_V387L_R (Table 3.1). Sequences were confirmed by Sanger Sequencing at UMGC with the primer set pBS_SK_M13_F and pBS_SK_M13_R (Table 3.1). This process of site-directed mutagenesis is illustrated in Supplemental Figure 3.2.

Subsequently, cDNA inserts of *BdSTP13* and its site-directed mutants were digested from pBS constructs at *EcoR*I-*Sal*I sites, separated on a 1% agarose gel in 1X TBE

and purified using a QIAquick® Gel extraction kit (Qiagen). The resulting fragments *BdSTP13*, *BdSTP13*^{G144R}, *BdSTP13*^{V387L} were each ligated at *Eco*RI-*Sal*I sites to the yeast expression vector backbone pDR196. For the co-expression analysis, *BdSTP13*$^{G144R+V387L}$ was inserted into both the pDR196 and pDR196T vector backbone. Each ligation product was transformed into *E. coli*, and the transformant was selected on 100 mg/ml ampicillin LB-agar plates after overnight incubation at 37°C. Single colonies were purified as described above. Sequences were verified by Sanger sequencing at UMGC using the primers MF127, MF128, pDR196_fPMA5 and pDR196_rADHclose (Table 3.1).

Functional analysis of hexose transport activity in yeast

Heterologous gene expression for the assessment of hexose uptake ability was performed as described by Milne et al (2017). The yeast (*Saccharomyces cerevisiae*) strain EBY.VW4000 (Wieczorke et al., 1999) is deficient in hexose transport and was used in these experiments. Vector constructs (pDR196-*BdSTP13*, pDR196-*BdSTP13*^{G144R}, pDR196-*BdSTP13*^{V387L}, pDR196-*BdSTP13*$^{G144R+V387L}$, pDR196T-*BdSTP13*$^{G144R+V387L}$) were each transformed into yeast using a PEG1000 transformation procedure (Dohmen et al., 1991). Media recipes for yeast growth and selection protocols are described by Milne et al. (2017).

The [^{14}C]hexose uptake assay was conducted following a previously established protocol (Milne et al., 2017) with the following modifications. Culture of a single colony of yeast was incubated overnight (220 rpm, 28°C) in 5 ml of liquid media (ura$^-$, trp$^-$ or both)

as a starter culture until the OD_{600} = ~1.0. Solid and liquid medium lacking uracil (ura^-) or lacking tryptophan (trp^-) were used for selection of pDR196 and pDR196T, respectively. For co-expression analysis, plasmids were co-transformed in equal quantity and selected on ura^-trp^- media (Milne et al., 2019). To prepare yeast culture for the uptake assay, 100 - 350 µl of starter culture were inoculated into 20 ml of liquid media and incubated overnight (220 rpm, 28°C) to OD_{600} = 0.6 - 1.0. Yeast cultures were harvested by centrifugation at 3,220 *g*, washed twice and re-suspended to a final OD_{600} of 2.0 with 25 mM MES-HEPES. Data were collected from three biological replicates, each representing a liquid culture from a single colony of yeast. For each biological replicate of the uptake assay, yeast cells were aliquoted into two technical replicates of 100 µl in 50 mL screw cap tubes.

Sugar uptake was assayed by measuring the retention of radiolabeled hexose in yeast cells after incubation with the uptake solution of 100 µM of [^{14}C] glucose or 1 mM of [^{14}C] fructose (pH = 5). Before the addition of 15 µl of uptake solution, cells were incubated in 15 µl of 100 mM ethanol for 1 min. At 4 min after exposure to the uptake solution, 100 µl of yeast cells were harvested with the funnel of a filter flask apparatus (Milne et al., 2017). A time-course experiment was also set up to assay the hexose transport capacity of BdSTP13; here yeast cultures were harvested at time intervals of 2, 4, 6, 8 min after exposure to the uptake solution. The radioactivity of the samples was measured using a Liquid Scintillation Analyzer for at least 2 min (PerkinElmer, Tri-Carb® 2810 TR), converted to units of µmol/g fw for glucose uptake or fructose uptake (g fw: gram fresh weight) (Milne et al., 2017). Empty vectors pDR196 or pDR196T were also included to test the basal hexose uptake of EBY.VW4000. Transport rates were calculated after

subtracting the background (empty vector control) uptake values, except for the time course experiment and the pH-dependence assays. To determine Michaelis-Menten Constant (Km) for glucose and fructose uptake, various concentration of [^{14}C] glucose (10, 20, 50, 100, 250, 500 µM) or [^{14}C] fructose (0.1, 0.2, 0.5, 1, 2.5, 5 mM) were used. A kinetic analysis of sugar transport rate (unit: µmol/g fw min^{-1}) (Milne et al., 2017) was conducted in GraphPad Prism 7 software (GraphPad Software, San Diego, CA, U.S.A.) using the Michaelis-Menten equation with least squares fit.

For pH-dependence assays, yeast cultures were buffered with 25 mM MES-HEPES with specified pH (pH = 4, 5, 6, 7, 8), and glucose uptake assays conducted as described above. For inhibitor assays, 10 µl of inhibitor solution was added 30 seconds before the addition of 100 mM of [^{14}C] glucose. Inhibitors were dissolved in DMSO. The final concentrations of the inhibitors were 100 mM for 2,4-dinitrophenol (DNP), 1.5 mM for Diethylpyrocarbonate (DEPC), 2 mM for N-ethylmaleimide (NEM) and 2 mM for phlorizin (Milne et al., 2017).

Results

Development of a screening pipeline to detect compatibility/incompatibility-associated genes in *Brachypodium distachyon* against rust fungi

The range of disease infection outcomes observed when Bd1-1 was challenged with three *P. graminis* formae speciales (*Pgt*, *Pg-pp*, or *Pg-lo*) created an experimental scenario to investigate the molecular and genetic events that confer NHR or partial resistance

phenotypes (Figueroa et al., 2013). A follow-up RNA-seq analysis was carried out by Gordon et al., 2017 and Melania Figueroa (unpublished data) to determine which genes were most affected by the rust infection in early infection time points (12 hpi and 18 hpi). The list of differentially expressed genes (DEGs, either up- or down-regulated) was ranked by expression changes relative to the mock infection control (Gordon et al., 2017). From here, a subset of 33 genes with more than two-fold changes for 'time point x isolate combination' (Supplemental Table 3.2) were selected. Hence, a screening pipeline was developed to determine if these genes were involved in resistance or susceptibility of *B. distachyon* to *Pgt* (Figure 3.1).

To establish this screening pipeline, a small collection of T-DNA insertion mutants in the *B. distachyon* Bd21-3 genetic background were requested from the Joint Genome Institute (Bragg et al., 2012; Hsia et al., 2017). This T-DNA mutant library has been characterized by Illumina or Sanger sequencing to define sequences flanking each T-DNA insertion site (Bragg et al., 2012; Hsia et al., 2017), making it possible to identify insertional mutants of genes of interest for functional analysis. The T-DNA inserts could either lead to a gene knockout mutant, or an overexpression mutant if the T-DNA containing a 4x35S enhancer element was inserted in the promoter region (Bragg et al., 2012). The T-DNA insertion mutants can be viewed in JBrowse of Phytozome 12 with the v3.1 annotation of *B. distachyon* (http://phytozome.jgi.doe.gov). I selected mutants that contained a predicted insertion within the gene model or within 500 bp of the promoter region of DEGs. In total, 34 T-DNA insertional sequence-indexed mutants representing 33

genes of interest (Supplemental Table 3.1), including two T-DNA insertion lines for *Bradi1g69710*, were requested for analysis and fed into the screening pipeline (Figure 3.1).

After setting up T_1 single-seed descent increases for each mutant, changes in resistance or susceptibility phenotypes compared to Bd21-3 (WT) after inoculation with *Pgt* W1101 were tested at T_2 generations for most mutant lines except for lines JJ9087, JJ13504, JJ10560, JJ3115, JJ19454, JJ661, CRC551, JJ8248, which I tested using the T_3 generation. As a measure of disease severity, the number of lesions (both necrotic and chlorotic) was recorded at 14 dpi. Overall, there was a high level of variation in this measure both within inoculation trials and also between trials, even for WT line Bd21-3, which complicated interpretation of the role of these genes in the interaction between *B. distachyon* and *Pgt* (data not shown). This was also complicated by the fact that some of the T_2 (or T_3) families were still segregating for the T-DNA insertion. Nevertheless, four T-DNA insertion mutants (JJ16783, JJ19596, JJ22746 and JJ21069, all tested within one T_2 family) showed an increased number of lesions relative to WT in at least two inoculation trials (Figure 3.2). These genes were predicted to encode a hexose transporter (*Bradi1g69710*), a WRKY transcription factor (*Bradi2g22440*), a cysteine-rich receptor-like kinase (*Bradi1g25552*), and a subtilisin-like protease (*Bradi3g10037*) (Table 3.2). Given the potential role of WRKY transcription factors in plant defense (Zheng et al., 2006) and hexose transporters in rust susceptibility (Moore et al., 2015), I focused on characterizing the insertional events in candidate genes *Bradi2g22440* and *Bradi1g69710* in more detail as described below.

A mutation in the WRKY transcription factor *BdWRKY46* gene resulted in enhanced susceptibility to *Pgt*

The gene *Bradi2g22440* is annotated as *BdWRKY46*, which falls into Group I of the WRKY transcription factor family in *B. distachyon* (Tripathi et al., 2012). According to the v.3.1 annotation of *Brachypodium* (http://phytozome.jgi.doe.gov), the closest ortholog of *BdWRKY46* in *Arabidopsis thaliana* was *AT2G38470* which encodes a WRKY transcription factor, *AtWRKY33* (Bragg et al., 2012; Hsia et al., 2017). *AtWRKY33* has been shown to play a role in phytoalexin biosynthesis upon activation of the mitogen-activated protein kinase (MAPK)-induced pathway (Zheng et al., 2006; Mao et al., 2011). Therefore, I utilized the T-DNA insertion line of *BdWKRY46* to characterize the role of the gene in defense responses against rust fungi.

T-DNA genotyping in BdWRKY46 mutants

The genome coordinates of *BdWRKY46* are 19,954,310 - 19,957,486 on chromosome 2 (referred to as Bd2) of the *B. distachyon* genome assembly v3.1 (International Brachypodium Initiative, 2010). The T-DNA insertion line JJ19596 (Bragg et al., 2012) is the product of the insertion of vector pJJ2LBA in the second intron of *BdWRKY46* at the site 19,955,057 on the Bd2 chromosome (Figure 3.3a). This line contains a single T-DNA insertion in its genome.

To characterize the role of *BdWRKY46* in *Brachypodium*-rust interactions, I requested additional T_1 seeds of JJ19596 to identify homozygous lines for the T-DNA

insertion of JJ19596. Two of the T_1 seeds of this line failed to germinate, and four of T_1 plants were viable for downstream genotyping analysis. Four individuals were genotyped using PCR to identify the presence of T-DNA inserts and identify homozygosity using the WT line Bd21-3 as a control. As shown in Figure 3.3b, multiplex PCR analysis genotyped #1 and #2 as wild type (WT1 and WT2) as they only displayed the wild-type PCR amplicon (1388 bp by MF86, MF87 primers) of *BdWRKY46.* This same approach identified one homozygous line (HM3) as it only showed the T-DNA band (~700 bp by MF87 and LB primer) which contained the region spanning the T-DNA left border and the genomic region. In addition, #4 was assigned as a heterozygous T-DNA insertion line (HET4), because it produced PCR products corresponding to both the wild type and the insertion alleles. Alignment of the sequences of PCR products to the reference genome of Bd21-3 validated the unique amplification of *BdWRKY46* and the presence of the T-DNA insertion. The results also suggested that the T_1 generation was segregating for this T-DNA insertion allele. The homozygous line JJ19596 HM3 was selfed to continue with the evaluation of disease phenotypes.

Disease phenotyping in homozygous BdWRKY46 mutants

The number of macroscopic lesions and fungal growth using a microscopic assay was evaluated in the *BdWRKY46* mutant, JJ19596, and WT Bd21-3 after infection with *Pgt* (isolate W1101 and 7a) at 14 dpi. T_2 progeny of the mutant JJ19596 HM3 (homozygous for the T-DNA insertion) were phenotyped. In response to both *Pgt* isolates, the total number of lesions in the T-DNA homozygous mutant plants was significantly higher than

the WT line (Mann-Whitney U test, *p* value = 0.001933 for W1101; *p* value = 0.03792 for 7a) (Figure 3.4a). Consistent with this, microscopic assays of the percentage of colonization (pCOL) and pustule formation (pPUST) also indicated that fungal growth was significantly greater in the homozygous mutant than in the WT line (Mann-Whitney U test, *p* value = 0.03571 for pCOL; *p* value = 0.03577 for pPUST) infected with *Pgt* W1101 (Figure 3.4b, c). Higher pCOL and pPUST values were also observed in the mutant after *Pgt* 7a infection, although this difference was not statistically significant. Overall, the disease phenotype patterns suggest that the homozygous T-DNA insertion mutant JJ19596 is more susceptible to stem rust infection than the WT line, and that *BdWRKY46* contributes to the NHR resistance of *B. distachyon* to *Pgt*. In addition, these data also indicate that disease outcomes in *B. distachyon* may be affected by the genotype of the rust isolate used for infection.

Characterization of sugar transporter *BdSTP13* in disease resistance against rust infection

The hexose transporter gene *Lr67* in wheat confers adult plant resistance (APR) against the three wheat rusts and powdery mildew (Moore et al., 2015). Sequence changes in *Lr67* (*Lr67res*) from the susceptibility allele (*Lr67sus*) on the wheat D genome lead to a loss in hexose transport ability and confer disease resistance (Moore et al., 2015). The candidate gene *Bradi1g69710* in *B. distachyon* is an ortholog of *Lr67* and belongs to the sugar transporter protein 13 (STP13) family. Given this, *Bradi1g69710* is referred to as

BdSTP13. Below I use the line JJ16783 to determine the role of *BdSTP13* in the interaction of *B. distachyon* and rust fungi.

Phylogenetic relationship of BdSTP13 with other hexose transporter members in Arabidopsis and other species

To understand the diversity and redundancy of STPs in *B. distachyon*, I conducted an orthology search using the fourteen members of the *Arabidopsis* STP family (Büttner, 2010) against the *B. distachyon* genome. This identified *Bradi1g69710* (*BdSTP13*) as the only orthologous gene for *AtSTP13* (Figure 3.5). *Bradi1g54920*, *Bradi5g12030*, *Bradi4g30290*, *Bradi1g54320*, *Bradi4g28000* were also identified as orthologs of *AtSTP1*, *5*, *7*, *8*, *14*, respectively. Phylogenetic inference based on the maximum-likelihood model indicated that *BdSTP13* forms a tight clade with wheat resistance gene *Lr67res* (Ta_Lr67_ALL26331) and its homeologs on 4A (Ta_4A_ALL26328), 4B (Ta_4B_ALL26329), and 4D (Ta_4D_ALL26330, namely *Lr67sus*), as well as barley HvSTP13 and STP13-like genes in the wheat ancestors *Triticum urartu* (*Tu*) and *Aegilops tauschii* (*Ae*) (Figure 3.5). Protein sequence identity between Lr67sus and BdSTP13 is 93.95% (Figure 3.6). In addition, alignment of STP13 proteins in several other plant species indicated that all STP13 members contain the two widely conserved amino acids Gly144 and Val387, except the wheat Lr67res allele (Arg144 and Leu387) (Figure 3.6) (Moore et al., 2015). The two amino acid changes of Lr67res were also not encoded in the genomes of 52 other natural inbred lines of *B. distachyon* (ABR2, ABR3, ABR4, ABR5, ABR6, ABR7, ABR9, Adi-10, Adi-12, Adi-2, Arn1, Bd1-1, Bd18-1, Bd2-3, Bd21-3, Bd21, Bd29-

1, Bd30-1, BdTR10C, BdRTR11A, BdTR11G, BdTR11I, BdTR12c, BdTR13C, BdTR13a, BdTR1i, BdTR2B, BdTR2G, BdTR3C, BdTR5I, BdTR7a, BdTR8i, BdTR9K, Bis-1, Foz1, Gaz-8, Jer1, Kah-1, Kah-5, Koz-1, Koz-3, Luc1, Mig3, Mon3, Mur1, Per1, RON2, S8iiC, Sig2, Tek-2, Tek-4, Uni2) (Gordon et al., 2014) (data not shown) suggesting that the resistance version of *Lr67* has not evolved in *B. distachyo*n.

BdSTP13 was induced after Pgt or Pca infection

Previous work on the STP13 family members reported that the expression of *STP13* in *Arabidopsis thaliana* and grapevine is induced by both biotic and abiotic stresses (Nørholm et al., 2006; Hayes et al., 2010; Lemonnier et al., 2014). Up-regulation of both *Lr67res* and *Lr67sus* alleles as well as their homeologs was also observed in wheat upon rust infection (Moore et al., 2015), suggesting that *Lr67* may be involved in plant responses to rust fungi. Therefore, I first investigated the transcript level of *Lr67sus* in the wheat McNair cultivar in a time course from 1 to 13 dpi after challenge with *Pgt* 7a (Figure 3.7a), using the wheat *GAPDH* (EU022331) (Yin et al., 2009) as reference gene for normalization. Statistical analysis was performed using one-way ANOVA for each time-course experiment followed by multiple comparisons to compare the gene expression changes between infected plants and mock inoculation. *Lr67sus* was significantly up-regulated at 9, 11, 13 dpi compared to mock treatment, when fungal sporulation was visible on wheat leaves. The highest levels of gene expression were detected at 9 and 11 dpi, with 4.7 ± 1.0 and 4.5 ± 0.8-fold change relative to mock treatment, respectively (Figure 3.7a).

To test if *BdSTP13* is also a pathogen-responsive gene like *Lr67sus*, I performed a similar time-course analysis in WT Bd21-3 to assay the expression level of *BdSTP13*, using the *Brachypodium GAPDH* as a reference gene for normalization (Figure 3.7b). After *Pgt* 7a challenge, *BdSTP13* was significantly induced at 3, 5, 7, 9 dpi compared to mock inoculations, with relative gene expression fold changes ranging from 2.0 ± 0.05 to 3.0 ± 0.4. The highest gene expression change was at 7 dpi. I also examined *BdSTP13* expression after inoculation with *Pca* isolate 12SD80 which produces more extensive colonization on WT Bd21-3 than *Pgt*. Inoculation with *Pca* 12SD80 also triggered a significant upregulation of *BdSTP13* expression at all time points (2.7 ± 0.04 to 8.4 ± 0.7-fold change), showing a generally higher gene induction level than after inoculation with *Pgt* 7a (Figure 3.7c). The highest expression level was detected at 9 and 11dpi, with both time points reaching nearly 8.5-fold change relative to the mock inoculation. These expression patterns suggest that the transcriptional regulation of *BdSTP13* is associated with rust infection in *Brachypodium*.

A mutation in BdSTP13 gene resulted in enhanced susceptibility to Pgt and Pca

T-DNA phenotyping in BdSTP13 mutants

The gene *BdSTP13* (*Bradi1g69710*) is localized at the sites of 68,416,021 - 68,421,356 on chromosome 1 (referred to as Bd1) of the genome reference of *B. distachyon* (accession Bd21, JGI v3.1 assembly) (International Brachypodium Initiative, 2010). To determine the role of *BdSTP13* in the *Brachypodium*-rust interaction, I first examined the

location of the T-DNA insert in the *B. distachyon* mutant JJ16783. In this line the T-DNA construct (pJJ2LBA, Bragg et al., 2012) was inserted within the fourth exon of *BdSTP13* at position 68,420,553 on chromosome Bd1 (Figure 3.8a). According to the *B. distachyon* T-DNA database, there are another three T-DNA inserts present in mutant JJ16783 (Bragg et al., 2012; Hsia et al., 2017). One of the T-DNA fragments was inserted at position 50,051,328 of chromosome Bd2 which corresponds to an intergenic region. Another T-DNA insertion occurred at position 7,766,785 on chromosome Bd1, ~84 bp before the start of a gene (*Bradi1g10720*) with predicted GTPase activity. Another T-DNA was predicted inserted in the intron of a gene (*Bradi3g08060*) encoding histone deacetylase 1. Thus, none of these three insertions are predicted to affect plant defense responses.

To identify homozygous T-DNA insertion mutants of *BdSTP13*, the T_1 progeny of JJ16783 was genotyped using a PCR-based approach similar to that described for the *BdWRKY46* mutant. Based on this analysis, two individuals (#1 and #10) displayed a WT genotype (WT1 and WT10) as they produced the gene-specific PCR amplicon (993 bp MF119 and MF120 primers), characteristic of the wild-type *BdSTP13* (Figure 3.8b). Three individuals (#2, #5, #8) were genotyped as homozygous for the T-DNA insertion as these individuals (HM2, HM5, HM8) exhibited no gene-specific products but instead showed two amplicons corresponding to the T-DNA flanking region (~819 bp by MF119 and LB primers; ~392 bp by MF120 and LB primers) (Figure 3.8b). The amplification of two different chimeric products in the homozygous mutant lines indicated the presence of tandem T-DNA inserts in a head-to-head (right-border-to-right-border) as depicted in Figure 3.8a (Tzfira et al., 2004) orientation at the targeted DNA site (Figure 3.8b and d).

Sanger sequencing of PCR products and alignment to the T-DNA LB sequence and the genomic flanking sequences confirmed the position and orientation of the T-DNA inserts in *BdSTP13*. Four T_1 individuals (#4, #6, #7 and #9) harbored both the gene-specific product and the two T-DNA flanking regions, and thus they were heterozygous for the T-DNA insertion (Figure 3.8c, d).

Disease phenotyping in homozygous BdSTP13 mutants

Based on the genotyping analysis, the T_1 homozygous mutants (JJ16783-HM2, HM5, HM8) were selfed to obtain T_2 progeny, which were evaluated for disease resistance phenotypes in a comparison to WT Bd21-3. Symptom development was evaluated at 14 dpi in response to *Pgt* 7a and *Pca* 12SD80 infection. At a macroscopic level, all three homozygous lines showed enhanced disease symptoms compared to the WT for both rust species, and sporulation was even observed after infection with *Pgt* 7a (Figure 3.9a). Homozygous mutant lines HM2, HM5, HM8 of JJ16783 showed a significantly higher number of lesions than WT Bd21-3 upon *Pgt* 7a treatment (statistics using linear mixed-effects model (lme4); p = 0.0017, 0.0164, 0.0445 for HM2, HM5, HM8, respectively) (Figure 3.9b). However, the extensive development of spreading chlorosis and necrosis prevented quantification of lesion numbers in the *Pca* infected plants (Figure 3.9a).

Gene expression level of BdSTP13 in homozygous mutants

To determine if the tandem T-DNA insertions affected the expression of *BdSTP13* in the homozygous mutant lines, gene expression was evaluated by reverse-transcription PCR (RT-PCR) in mock-inoculated and rust-treated (*Pgt* 7a and *Pca* 12SD80) plants at 4 dpi. In this experiment, I used PCR primers (positions shown in Figure 3.8a) directed to the *BdST13* transcript, either upstream of the T-DNA insertion site (primers MF125 and MF126) or spanning the T-DNA insertion site (MF129 and MF130). The *GAPDH* gene in *B. distachyon* was used as a reference gene. According to this analysis, all three homozygous mutants displayed the presence of the transcript upstream of the T-DNA insertion (Figure 3.10). Consistent with the gene expression profile *(*Figure 3.7b, c), levels of this part of transcript of *BdSTP13* were higher in both WT and mutant samples treated with *Pgt* 7a or *Pca* 12SD80 than mock inoculation. However, the transcript region spanning the T-DNA insertion was only detected in the WT line and was absent in the homozygous mutant lines (Figure 3.10), indicating that the mutation abolished the production of an intact transcript, and therefore it is unlikely that the BdSTP13 protein is functional in the homozygous mutants.

Functional validation of BdSTP13 in heterologous gene expression system in yeast

Time course of hexose uptake of BdSTP13

The hexose transport activity of *BdSTP13* was determined by employing a heterologous expression system in the yeast (*Saccharomyces cerevisiae*) strain

EBY.VW4000, which is deficient in hexose transport (Wieczorke et al., 1999). Yeast cells expressing the empty vector pDR196 as a negative control showed no specific uptake of [^{14}C] glucose or [^{14}C] fructose, with background accumulation in the range of 0.01 - 0.02 µmol/ g fw over the 8 min time course (Figure 3.11). However, yeast transformed with *BdSTP13* (with vector pDR196) showed a linear increase in accumulation of both glucose and fructose, at a rate of 2 and 0.75 µmol/ g fw per minute (Figure 3.11). In a similar experiment, the rate of glucose uptake into yeast cells expressing the wheat *Lr67sus* was previously measured as 1 - 1.5 nmol/g fw per minute (Moore et al., 2015). For the subsequent analysis, yeast cells were harvested at 4 min after incubation with hexose uptake solution.

Transport affinity of glucose and fructose of BdSTP13

To determine the affinity of the sugar transporter BdSTP13 for different hexose substrates, kinetic analyses of uptake were performed using various concentrations of glucose and fructose (Figure 3.12). The results demonstrated that hexose transport by BdSTP13 follows the Michaelis–Menten kinetics model and is concentration dependent. At low glucose or fructose concentration, the uptake activity increased steeply with increasing substrate concentration, whereas at high concentrations the transporter was saturated and the rate of transport reached a stable plateau. The kinetic parameter *Km* (Michaelis constant) represents the concentration at which the transport activity is 50% of its maximum (saturated) rate and indicates the affinity of a transporter to its substrate, with low *Km* values indicating high affinity (Crowley, 1975). BdSTP13 displays a higher

affinity for glucose uptake (*Km* = 104 ± 7.05 μM) compared to fructose (*Km* = 430 ± 20 μM) (Figure 3.12). These values are similar to those for wheat Lr67sus, which had a *Km* value of 98.8 μM and 1.32 mM for glucose and fructose transport, respectively (Moore et al., 2015).

BdSTP13 is a proton-hexose symporter

Most homologous STP13 family members are described as hexose-proton symporters which require a proton gradient for hexose transport (Slewinski, 2011; Moore et al., 2015). To test if the ability of hexose uptake by BdSTP13 is pH-dependent, I conducted a glucose uptake assay in the *BdSTP13*-expressing yeast EBY.VW4000 using glucose uptake solutions with pH ranging from 5 to 8. As indicated in Figure 3.13, the glucose transport rate of BdSTP13 is pH dependent, with the highest and lowest capability at pH = 5 and pH = 8, respectively. The proton ionophore 2,4-dinitrophenol (DNP) dissipates the proton gradient across membranes (Gear et al., 2000; Moore et al., 2015) and strongly inhibited glucose uptake of BdSTP13 to ~20% of the uninhibited rate (Figure 3.14). Treatment of yeast cells with diethyl pyrocarbonate (DEPC) or N-ethylmaleimide (NEM), which disrupts energy metabolism (Gear et al., 2000) also reduced glucose uptake, with inhibition ranging from 33% and 45%, respectively. These compounds modify amino acid side-chains of proteins and can disrupt energy metabolism in cells uptake (Felle et al., 1983; Lemoine and Delrot, 1987; Gear et al., 2000). In addition, a non-transported glucose analog, phlorizin, which competes with glucose binding (Gear et al., 2000; Moore et al.,

2015), also decreased to the transport rate of BdSTP13, confirming its role as a specific hexose transporter.

BdSTP13 site-directed mutants share similar transport activity as those in Lr67sus

As discussed above, BdSTP13 shares the two key amino acids (Gly144 and Val387) with Lr67sus that were mutated in Lr67res and led to a loss of hexose transport ability. Using site-directed mutagenesis (SDM), I recapitulated these *Lr67res*-specific mutations in *BdSTP13* either singly in $BdSTP13^{G144R}$ and $BdSTP13^{V387L}$ or combined in $BdSTP13^{G144R+V387L}$. Heterologous expression of these mutants in yeast EBY.VW4000 showed that introducing these mutations reduced the hexose transport capacity of *BdSTP13* (Figure 3.15). Furthermore, the single mutant $BdSTP13^{G144R}$ and the double mutant $BdSTP13^{G144R+V387L}$ fully abolished the ability of hexose uptake while the $BdSTP13^{V387L}$ mutant reduced the transport rate by ~50%, compared to *BdSTP13* as a positive control. This finding indicated that the G144R mutation was primarily responsible for the difference in transport activity, as was determined for Lr67res (Moore et al., 2015). The two amino acid polymorphisms (G144R, V387L) in Lr67res have a dominant-negative effect on hexose transport, and therefore I also evaluated the hexose transport rate in a yeast co-expressing *BdSTP13* and $BdSTP13^{G144R+V387}$ (Figure 3.16). Here, the yeast transformant with the backbone vector pDR196 or pDR196T were used as negative controls. Data from this experiment indicated that co-expression of *BdSTP13* and $BdSTP13^{G144R+V387L}$ led to ~50% reduction of glucose uptake rate compared to the wild-type allele of *BdSTP13*. In

summary these functional analyses indicate *BdSTP13* shares similar hexose transport function with *Lr67sus*.

Discussion

In recent years, extensive open genetic resources have become available and enabled powerful studies of the biology of *B. distachyon* (International Brachypodium Initiative, 2010; Bragg et al., 2012; Gordon et al., 2017; Hsia et al., 2017). The collection of *Brachypodium* T-DNA insertion mutants generated by JGI (Bragg et al., 2012) is an example of these resources. Cross-reference of this collection of mutants with a list of genes of interest based on gene expression profiles upon *Pgt* infection allowed us to examine the role of 33 genes with defense-associated functions in contributing to susceptibility or resistance to wheat stem rust fungus (Supplemental Table 3.1, 3.2).

The infection of *B. distachyon* by the non-adapted pathogen *Pgt* is characterized by partial fungal growth (Figueroa et al., 2013, 2015), which complicates the interpretation of experiments that rely on comparison of macroscopic phenotypes. Thus, subtle rust infection differences between mutants and WT lines may be difficult to detect. As encountered in this study, the high variation of infection phenotypes within a single genotype of *B. distachyon* across experiments prevented in most cases robust conclusions as to whether or not particular genes play a role in NHR against *Pgt*. Such levels of variation may have been attributed to subtle changes in environmental conditions between experiments, especially the levels of humidity. In spite of these complications, this study

identified four genes (Table 3.2) that may contribute to resistance responses associated with NHR. Two of these, a WRKY transcription factor *BdWRY46* and a sugar transporter *BdSTP13*, were followed up with examination of homozygous T-DNA insertion lines. Data presented here suggests that both genes contribute to NHR of *B. distachyon* to cereal rusts.

BdWRKY46 in defense responses against rust fungi

WRKY transcription factors (WRKY TFs) are one of the best described transcription factor families involved in plant defense responses. WRKY TFs specifically bind to W-box elements containing a (T)TGAC(C) motif in the promoters of targeted genes (Eulgem et al., 2000). Network analysis identified co-expression of WRKY TFs with numerous W-box containing defense-related genes, suggesting a critical role of WRKY TFs in transcriptional regulation in plant immunity (Eulgem and Somssich, 2007; Somssich, 2007). Research on parsley demonstrated that three WRKY TFs are activated upon perception of fungal PAMPs and subsequently activate expression of pathogenesis-related (*PR*) genes as part of the PTI response (Rushton et al., 1996; Eulgem et al., 1999). Another study also showed that a WRKY TF, *AtWRKY33,* in Arabidopsis plays an essential role in PTI against the necrotrophic pathogen *Botrytis cinerea* (Mao et al., 2011; Birkenbihl et al., 2012). After pathogen- or PAMP-treatment (such as bacterial flagellin flg22), *AtWRKY33* is phosphorylated in a mitogen-activated protein (MAP) kinase cascade, and activates a massive transcriptional network of host responses, including those associated with phytoalexin biosynthesis (Mao et al., 2011). In addition, *AtWRKY33* acts as a transcriptional regulator of plant defense responses associated with the hormones jasmonic

acid (JA) and salicylic acid (SA). Previous work indicated that *AtWRKY33* positively modulates the JA-defense signaling pathway while negatively affecting SA-signaling, and therefore contributes to disease resistance against *B. cinerea* and the hemibiotrophic pathogen *Pseudomonas syringae* (Zheng et al., 2006). It is well studied that JA and SA have antagonist roles in responding to resistance against necrotrophic and biotrophic pathogens (Glazebrook, 2005; Zheng et al., 2006), because SA-induced cell death is critical for disease resistance against biotrophic pathogens but can promote the growth of necrotrophic pathogens (Govrin and Levine, 2000; Glazebrook, 2005; Li et al., 2006).

BdWRKY46 is an ortholog of *AtWRKY33* in Arabidopsis, and this work suggests that *BdWRKY46* may contribute to resistance of *B. distachyon* to the biotrophic pathogen *Pgt.* This is interesting because the role of *AtWRKY33* or its orthologs in plant immune responses against biotrophic pathogens is largely uncharacterized. The homozygous T-DNA insertion mutant JJ19596 allowed enhanced fungal growth compared to WT, as indicated by the larger number of lesions and the higher percentage of colonization and pustule formation (Figure 3.4). T-DNA insertions in the intron regions usually generate a knockout phenotype (Rosso et al., 2003; Hsing et al., 2007; Hsia et al., 2017); therefore future work to determine the effect of the T-DNA in the mutant using gene expression analyses should be carried out to evaluate if *BdWRKY46* produces an aberrant transcript. If *BdWRKY46* shared a similar gene regulatory function as *AtWRKY33*, loss of function of *BdWRKY46* would likely induce the expression of SA-responsive genes, and thereby facilitating cell death and inhibiting rust growth. However, the results shown here contradict this idea, since the homozygous T-DNA insertional mutant showed enhanced

rust growth. This may suggest a different function for *BdWRKY46* in *B. distachyon*. Considering this, another important follow-up experiment would be to overexpress *BdWRKY46* and evaluate whether this leads to enhanced or reduced disease phenotypes. The expression of defense-related genes such as marker genes in the SA and JA signaling pathways could also be investigated in the *BdWRKY46* mutant line or in overexpressing lines to characterize its role in plant defense. The hypothesis that *BdWRKY46* plays a regulatory role in hormone pathway crosstalk during PTI responses when sensing *Pgt* in *B. distachyon* still needs to be addressed.

Furthermore, the expression profiles of *BdWRKY46* at various time points could be examined in both WT and mutant to understand if this WRKY TF acts in early or later infection stages. It also remains to be determined whether modifications of the ortholog of this gene in wheat can also result in susceptibility in the wheat-rust pathosystem. Thus, the overexpression of the ortholog of this transcription factor in wheat could also be investigated. These aspects should be considered in the future as part of the functional characterization of *BdWRKY46* and its ortholog in wheat.

BdSTP13 in defense responses against rust fungi

It has been proposed that one important mechanism for NHR is the restriction of nutrient availability by the plant via sugar transporters to stop the growth of non-adapted pathogens particularly bacteria (Senthil-Kumar and Mysore, 2013). A class of sugar transporters that is well-studied is the *SWEET* family, which encodes sugar efflux

transporters that are targeted by bacterial pathogens to redirect nutrient flux to the plant apoplast for nutritional purposes (Chen et al., 2010). The sugar transporter 13 (STP13) family, known for their sugar influx transporter capacity (Büttner, 2010) have also been shown to be involved in plant disease responses (Moore et al., 2015; Dodds and Lagudah, 2016; Yamada et al., 2016; Milne et al., 2019). Research on *AtSTP13* indicates that this gene is induced under biotic and abiotic stresses and is associated with programmed cell death (Nørholm et al., 2006; Slewinski, 2011). In addition, Yamada et al (2016) demonstrated that *AtSTP13* activity is PAMP-inducible, and that the transporter physically interacts with the PRR FLS2 (flagellin receptor flagellin-sensitive 2) and the FSL2 co-receptor BAK1 (BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1), indicating a role of AtSTP13 in PTI. The phosphorylation of AtSTP13 by BAK1 was shown to be required for antimicrobial defense (Yamada et al., 2016).

The screening pipeline for disease phenotype changes in *B. distachyon* mutants against *Pgt* identified the sugar transporter gene *BdSTP13*, as a gene with a potential role in NHR. *BdSTP13* is an ortholog of a wheat multi-pathogen resistance gene *Lr67* (*Lr67res*) (Moore et al., 2015). The *Lr67res* allele is a recently evolved mutant derived from the *STP13* gene (*Lr67sus*) in wheat. *Lr67sus* encodes a hexose transporter and its function depends on two conserved amino acid residues (Gly144 and Val387). The amino acid specific mutations G144R and V387L in *Lr67res* lead to the abolishment of sugar transport capacity. As shown in this chapter the residues Gly144 and Val387 are highly conserved in all STP13-like family members in various species, including BdSTP13, which suggests that they are derived from an ancient common ancestor (Figure 3.5, Figure 3.6). Therefore,

BdSTP13 may share similar functions as *Lr67sus* in wheat and other *STP13* members from other species such as Arabidopsis *AtSTP13*.

Several hypotheses have been proposed to explain the underlying mechanism by which *Lr67* confers resistance against multiple biotrophic pathogens (Moore et al., 2015; Dodds and Lagudah, 2016; Milne et al., 2019). The first hypothesis is that the presence of Lr67res blocks hexose retrieval from the apoplast in infected leaves, thus reducing the nutrient availability to haustoria from within the cell and limiting fungal growth. An alternative explanation is that inhibition of hexose transport results in an enhanced hexose/sucrose ratio in the leaf apoplast and activates the sugar-associated defense signaling pathway (Herbers et al., 1996). A third hypothesis is that *Lr67res* represents a neo-functionalized allele with an unknown novel activity that induces disease resistance (Milne et al., 2019). One of the challenges in addressing these issues is the difficulty of producing a complete knockout of *Lr67* together with its homeologs in hexaploid wheat. *B. distachyon* as a close diploid relative of wheat, offers an opportunity to generate *Lr67*-like deficient knockouts.

This study shows that *BdSTP13* is a pathogen-responsive gene that is upregulated by both *Pgt* and *Pca* inoculations. In addition, *BdSTP13* expression in *Pca*-infected plants was more strongly induced in most time points than *Pgt*-inoculated plants, likely due to a higher compatibility between *B. distachyon* and *Pca*. Overall, time-course expression analysis of *BdSTP13* and *Lr67sus* suggests that these sugar transporters may play a more important role at a later stage of the infection when the pathogen undergoes spore differentiation (*Pgt* in wheat) or forms large colonies (*Pgt* or *Pca* in *B. distachyon*). In

addition, phenotypic results presented here showed that a T-DNA induced knockout of *BdSTP13* resulted in enhanced susceptibility to *Pgt* and *Pca*, as observed by the increased macroscopic symptoms with both pathogens (Figure 3.9a). A significantly increased lesion count was also observed in all three homozygous mutant lines (JJ16783-HM2, HM5, HM8) infected with *Pgt* (Figure 3.9b). These analyses suggest that *BdSTP13* acts as a resistance factor in *Brachypodium*-rust interactions. This hypothesis could be further tested by assaying rust colonization using the fungal biomass assay or by microscopic analysis of percentage of colonization and pustule formation in infected leaves of WT and homozygous mutant plants. In addition, segregation analysis of T_2 heterozygous mutants could be conducted to evaluate if an increased disease symptom phenotype segregates with T-DNA induced knockout mutants. Future research directions may also include expression of *BdSTP13* containing the Lr67res mutations in transgenic *B. distachyon* and using CRISPR/Cas9 to generate these changes or knockout mutations in the native *BdSTP13* gene.

Similar to other STP13 orthologs such as *AtSTP13, HvSTP13* and *Lr67sus* (Nørholm et al., 2006; Moore et al., 2015; Milne et al., 2019), heterologous expression of *BdSTP13* in yeast cells demonstrated that *BdSTP13* acts as a high-affinity to glucose, hexose/H^+ symporter (Figure 3.12, Figure 3.13, Figure 3.14). Treatment with the proton gradient disruptor DNP greatly inhibited glucose uptake in *BdSTP13*-expressing yeast cells, consistent with H+-driven uptake. The optimum pH 5 for the glucose transport capacity of *BdSTP13* is consistent with the acidic environment of the plant apoplast (pH ~5) (Yu et al., 2000). The acidic apoplastic environment provides the necessary proton gradient to drive

hexose transport into the plant cytoplasm which is typically at a pH of ~7.5 (Kurkdjian and Guern, 1989). Phlorizin, as a non-transported competitive inhibitor of hexose uptake (Felle et al., 1983; Lemoine and Delrot, 1987; Gear et al., 2000), also inhibited uptake of BdSTP13, confirming that it was mediated by substrate-specific mechanism. The amino acid modifying chemicals DEPC and NEM also inhibited uptake, which may be due to direct modification of the BdSTP13 protein or due to more general effects in disrupting other membrane proteins such as H+ transporters or energy metabolism (Bush, 1993; Gear et al., 2000).

It is difficult at this point to explain how Lr67res confers resistance to biotrophic pathogens like rust fungi. These assays discussed in this Chapter confirm BdSTP13 functions as a glucose transporter, so it is possible that the *BdSTP13*-knockout mutant lacks hexose transport ability *in planta*. The results presented here showed increased rust infection in a *BdSTP13*-knockout mutant, which contradicts the hypothesis that BdSTP13 acts as a susceptibility factor as was proposed for *Lr67sus* (Moore et al., 2015). It also negates the hypothesis that limiting pathogen nutrition to rust haustoria contributes to resistance phenotypes (Moore et al., 2015; Dodds and Lagudah, 2016). An alternative explanation is that disruption of the *BdSTP13* gene likely affects sugar-mediated defense signaling pathways in *B. distachyon,* resulting in susceptibility. Under this scenario, the loss of hexose transport ability of *BdSTP13* could lead to an increased accumulation of glucose and altering the hexose/sucrose ratio in the apoplast, and affecting responses to rust infection. For example, previous studies have shown that overexpression of invertases

results in glucose accumulation in the apoplast and induction of expression of defense-related genes such as pathogenesis-related (*PR*) genes (Herbers et al., 1996).

Overall, this chapter demonstrates that *B. distachyon* can be a suitable system to investigate the underlying mechanisms of NHR against rust fungi. The discovery of defense-related genes in *B. distachyon* may offer new resistance mechanisms for wheat disease resistance engineering given that *B. distachyon* and cereal crops share a close phylogenetic relationship. Thus, a genetic strategy incorporating diverse resistance mechanisms, such as in *B. distachyon*, to control rust diseases may provide opportunities to innovate and deliver solutions to boost wheat production by non-traditional breeding technologies and minimize the risk of pathogen evolution and durable and broad-spectrum disease resistance.

Table 3.1. Primers

Primer name	Primer sequence (5'-3')	References
***Brachypodium* T-DNA genotyping**		
MF86_BdWRKY46_F3	CCCTAGCCTTCCTCAACTCC	
MF87_BdWRKY46_R3	TGCTCCACCTTCTTCTTCGT	
MF119_ BdSTP13_F1	ACGGATAAATACTGAAACTGCGT	
MF120_ BdSTP13_R1	CGGCGTTGTTGATGGTGTT	
T3 T-DNA LB	AGCTGTTTCCTGTGTGAAATTG	Hsia et al., 2017
RT-qPCR and RT-PCR		
GAPDH_F	TTGCTCTCCAGAGCGATGAC	Hong et al., 2008
GAPDH_R	CTCCACGACATAATCGGCAC	
MF125_ BdSTP13_qF1	GCGGGCTCAACATCCTGT	
MF126_ BdSTP13_qR1	CACCGCCTTGCCTTCGTC	
MF129_ BdSTP13_qF2	CTGGTTTCGGTCTACTGCGT	
MF130_ BdSTP13_qR2	AGATGGCGTACTTGAGGTGG	
Lr67_DD F	CAGCACTGGTTCTGGAAGAGA	Moore et al., 2015
Lr67_DDR	GTTACATTTACATAGTAGGAGC	
MF142_TaGAPDH_F	TTGCTCTGAACGACCATTTC	modified from Yin et al, 2009
MF143_TaGAPDH_R	GACACCATCCACATTTATTCTTC	
Yeast expression vector construction for *BdSTP13*		
MF113_BdSTP13_EcoRIF	ATAGCGAATTCATGGCGGGCGGAGGGTT	
MF114_BdSTP13_SalIR	TATGCGTCGACTCAAACGGCGGCGTTGTTGAT	
pBS_SK_M13_F	GTAAAACGACGGCCAGT	addgene.org
pBS_SK_M13_R	GGAAACAGCTATGACCATG	
MF127_F	CTTCGCCAACCTCGTCAAC	
MF128_R	GAACGTCTCGCTGGGGAT	
pDR196_fPMA5	CTCTCTTTTATACACACATTC	addgene.org
pDR196_rADHclose	CGTAATACGACTCACTATAGG	
Site-directed mutagenesis		
BdSTP13_G144R_F	AGGATCCTGCTCCGCTGTGGCGTCG	
BdSTP13_G144R_R	CGACGCCACAGCGGAGCAGGATCCT	
BdSTP13_V387L_F	GGGTGGGCCGTATTGTTGGTGGTCATGGTG	
BdSTP13_V387L_R	CACCATGACCACCAACAATACGGCCCACCC	

***The primers were designed using the Primer3 software, except those specified with references.**

Table 3.2. Candidate genes involved in *Brachypodium distachyon* defense responses against rust fungi.

T-DNA mutant	Gene	Putative Function Description
JJ16783	*Bradi1g69710*	Sugar transporter
JJ19596	*Bradi2g22440*	WRKY Transcription Factor
JJ22746	*Bradi1g25552*	Cysteine-Rich Receptor like kinase
JJ21069	*Bradi3g10037*	Subtilisin-like protease (subtilase)

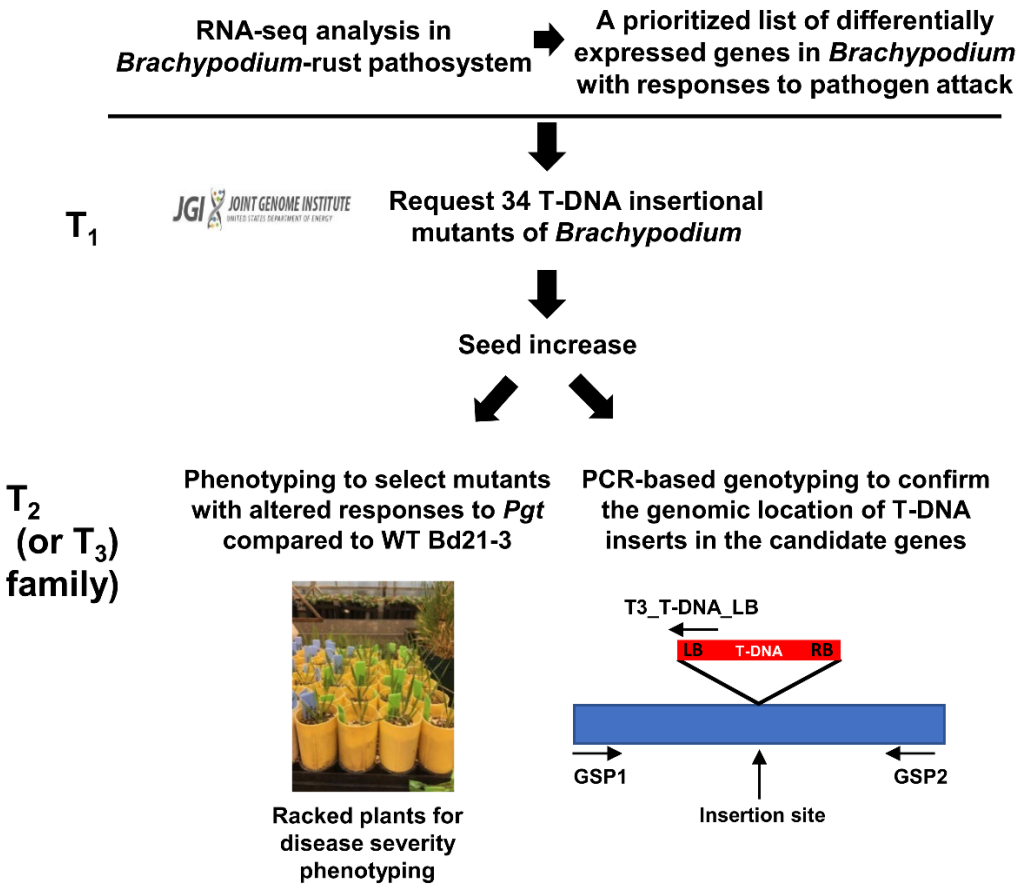


Figure 3.1. Workflow for disease phenotype screening T-DNA insertional mutants of *B. distachyon* by infection with *P. graminis* f. sp. *tritici*.
Diagram illustrates pipeline to identify *B. distachyon* genes associated with disease susceptibility or resistance using reverse genetics. The seeds for phenotyping and genotyping purpose were obtained via single-seed descent increases of T_1 progeny. Individuals from T_2 or T_3 families were inoculated with the *P. graminis* f. sp. *tritici* (*Pgt*) isolate W1101 and compared to WT Bd21-3 as a control. Schematic drawing (right side) shows a gene of interest and T-DNA insert (LB and RB denote left and right border, respectively) and location of primers used for PCR-based genotyping of mutant lines. GSP1 and GSP2 are gene-specific primers that target amplification of the wild-type allele. The primer T3_T-DNA_LB (Hsia et al., 2017) targets the T-DNA insertion allele, and combined with primer GSP1 or GSP2 allows amplification of T-DNA insert flanking regions depending on the orientation of the T-DNA fragment.

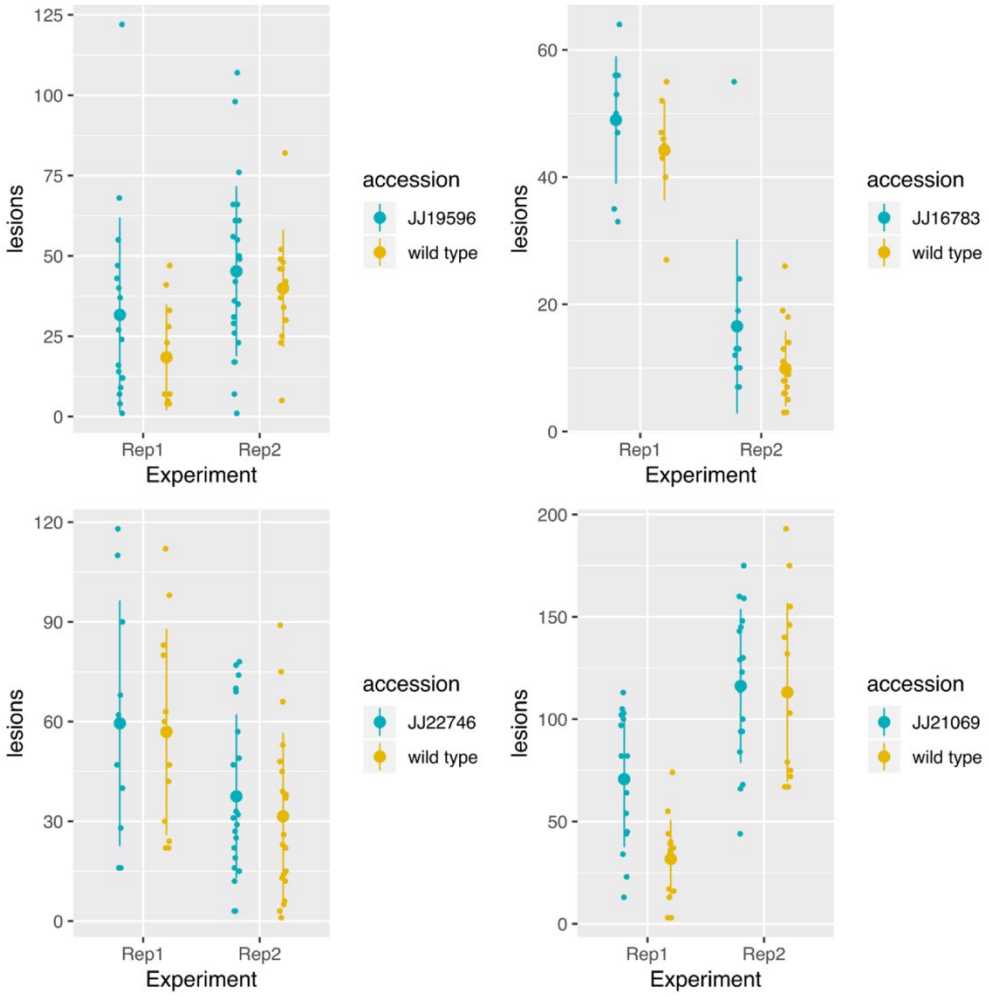


Figure 3.2. Quantification of macroscopic lesions in T-DNA insertional mutants of *B. distachyon* after inoculation with *P. graminis* f. sp. *tritici*.
Strip charts representing the number of lesions per individual of T-DNA insertional mutant (T_2 progeny, blue) or wild type (Bd21-3, yellow) after infection with *Pgt* W1101. Results from two independent experiments (replicates) are shown. Vertical line represents the mean (large dot) ± standard deviation (vertical line) in each independent experiment (replicate, Rep). *B. distachyon* line 19596 contains a T-DNA insertion in gene *Bradi2g22440* (*BdWRKY46*, WRKY transcription factor). JJ16783 contain a T-DNA insertion in gene *Bradi1g69710* (*BdSTP13*, sugar transporter). JJ22746 contains a T-DNA insertion in gene *Bradi1g25552* (cysteine-rich receptor like kinase). JJ21069 contains a T-DNA insertion in and *Bradi3g10037* (subtilisin-like protease).

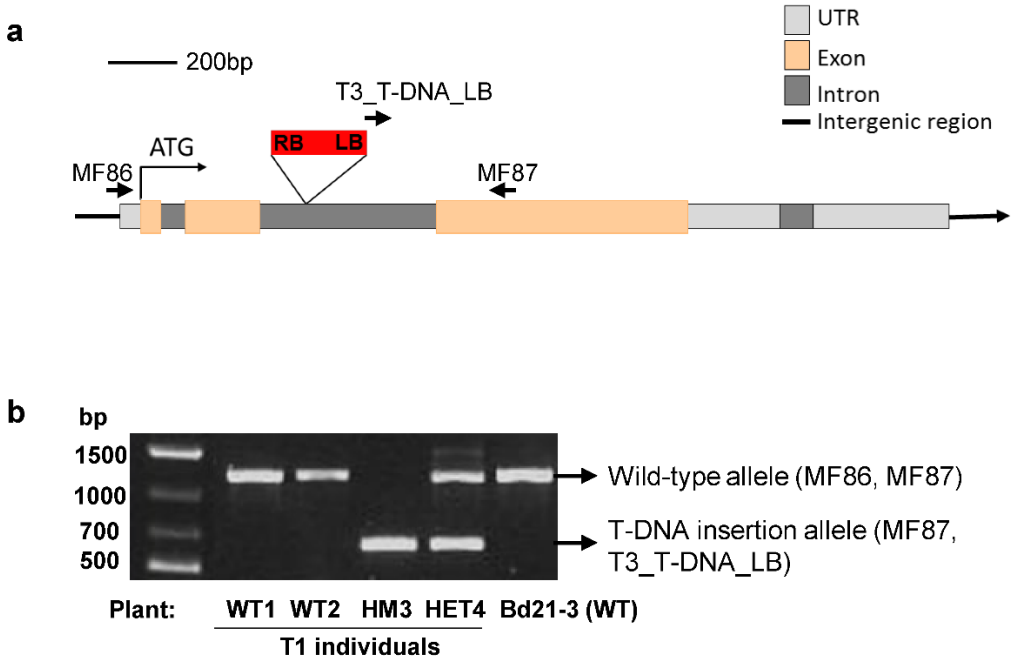


Figure 3.3. PCR-based T-DNA mutant genotyping of line JJ19596.
a) Diagram of the disruption of *BdWRKY46* gene (*Bradi2g22440*) by T-DNA mutagenesis. The boxed region indicates the gene transcript with the untranslated region (UTR) in light grey box, coding sequence in orange and introns in dark grey. Backbone arrow depicts 5'-3' orientation. The location of the start codon (ATG) is shown (chromosome Bd2: 19,954,490) above gene model. The T-DNA insertion site (chromosome Bd2: 19,955,057) is depicted by a red box above the second intron of *Bradi2g22440*. Short arrows illustrate primers used for PCR-based genotyping of T-DNA mutant. **b)** Multiplex PCR analyses for genotyping of T_1 individuals from line JJ19596 using gene-specific primers MF86, MF87 to target amplification of wild-type gene and the T-DNA primer T3_T-DNA_LB combined with primer MF87 to target flanking regions of the insert. A gel electrophoresis image with GeneRuler 1kb DNA ladder (shown in the first lane) shows PCR products. DNA template from wild type Bd21-3 was used as positive experimental control. Genotypes of T_1 individual plants are defined as WT to indicate homozygosity of the wild-type gene; HM to indicate homozygosity of disrupted gene; and HET to indicate heterozygosity of the wild-type gene as the disrupted gene is also present.

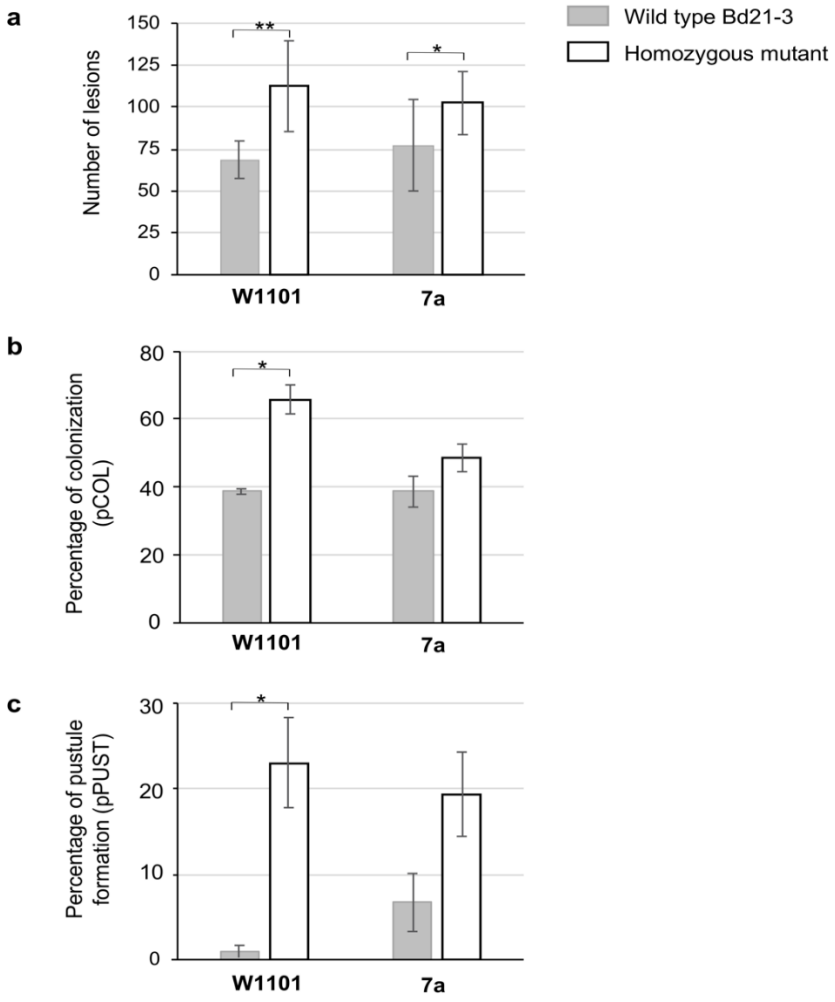


Figure 3.4. Phenotypic characterization of homozygous mutant lines JJ19596 after infection with *P. graminis* f. sp. *tritici*.

a) Bar graphs show mean (± standard deviation, SD) of lesion number in wild type Bd21-3 (grey bars) or T_2 homozygous mutant individuals from line JJ19596 (white bars) after infection with *Pgt* isolates W1101 and 7a. **b)** and **c)** Bar graphs show mean (± SD) of the percentage of *Pgt* colonization (pCOL) and percentage of pustule formation (pPUST), respectively after infection with *Pgt* isolates W1101 and 7a. pCOL and pPUST values were assayed from five secondary leaves stained with wheat germ agglutinin Alexa Fluor 488 conjugate (WGA-Alexa) for pathogen visualization under a fluorescence microscope (Leica). Data shown in **a)** to **c)** panels represents one independent experiment containing eight plants of wild type (Bd21-3) or homozygous mutant plants (HM4). Statistical analysis was performed using Mann-Whitney test. Two asterisks represent p value < 0.01 while one asterisk represents p value < 0.05.

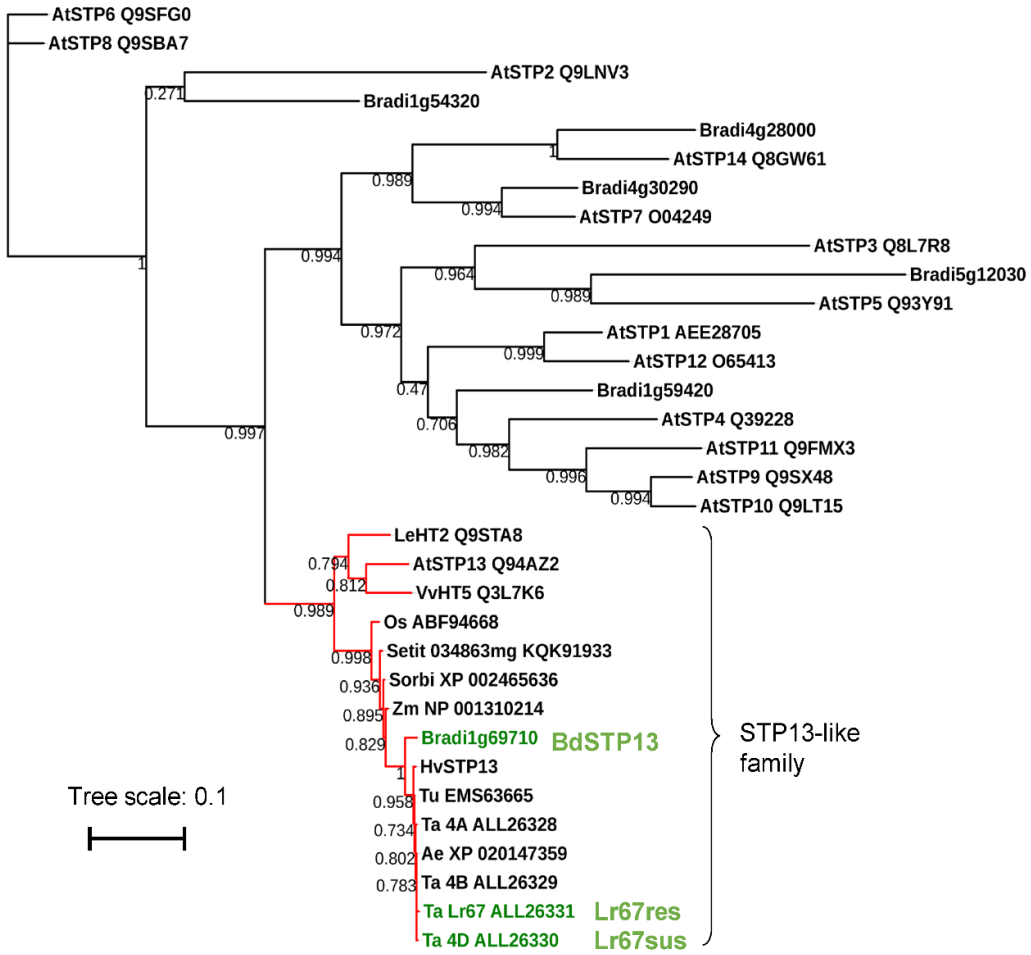


Figure 3.5. Phylogeny analysis of predicted sugar transporter genes in various plant species including *B. distachyon*

A maximum-likelihood-based phylogenetic tree built from protein sequence alignments, including sugar transporter proteins from *Arabidopsis thaliana* (AtSTP1_AEE28705, AtSTP2_Q9LNV3, AtSTP5_Q93Y91, AtSTP6_Q9SFG0, AtSTP7_O04249, AtSTP8_Q9SBA7, AtSTP9_Q9SX48, AtSTP10_Q9LT15, AtSTP11_Q9FMX3, AtSTP12_O65413, AtSTP14_Q8GW61), *B. distachyon* (Bradi1g54920, Bradi5g12030, Bradi4g30290, Bradi1g54320, Bradi4g28000, Bradi1g69710 (BdSTP13)), *Triticum aestivum* (Ta_4A_ALL26328, Ta_4D_ALL26330 (*Lr67sus* allele), Ta_4B_ALL26329, Ta_Lr67_ALL26331 (*Lr67res* allele), *Triticum urartu* (Tu_EMS63665), *Sorghum bicolor (*Sorbi_XP_002465636), *Aegilops tauschii* (Ae_XP_020147359), *Zea mays* (Zm_NP_001310214), *Oryza sativa (*Os_ABF94668), *Setaria italica* (Setit_034863mg), *Hordeum vulgare* (HvSTP13), *Vitis vinifera* (VvHT5_Q3L7K6) and *Lycopersicon esculentum* (LeHT2_Q9STA8). Clades shown in red define the STP13-like family. *Bradi1g69710* (BdSTP13), and wheat Ta_Lr67_ALL26331 (Lr67res allele) and Ta_4D_ALL26330 (Lr67sus allele) are indicated by green front. Scale bar represents a unit of branch length that corresponds to 0.1 substitutions per site. Bootstrap support values for branch bifurcations are shown next to each node.

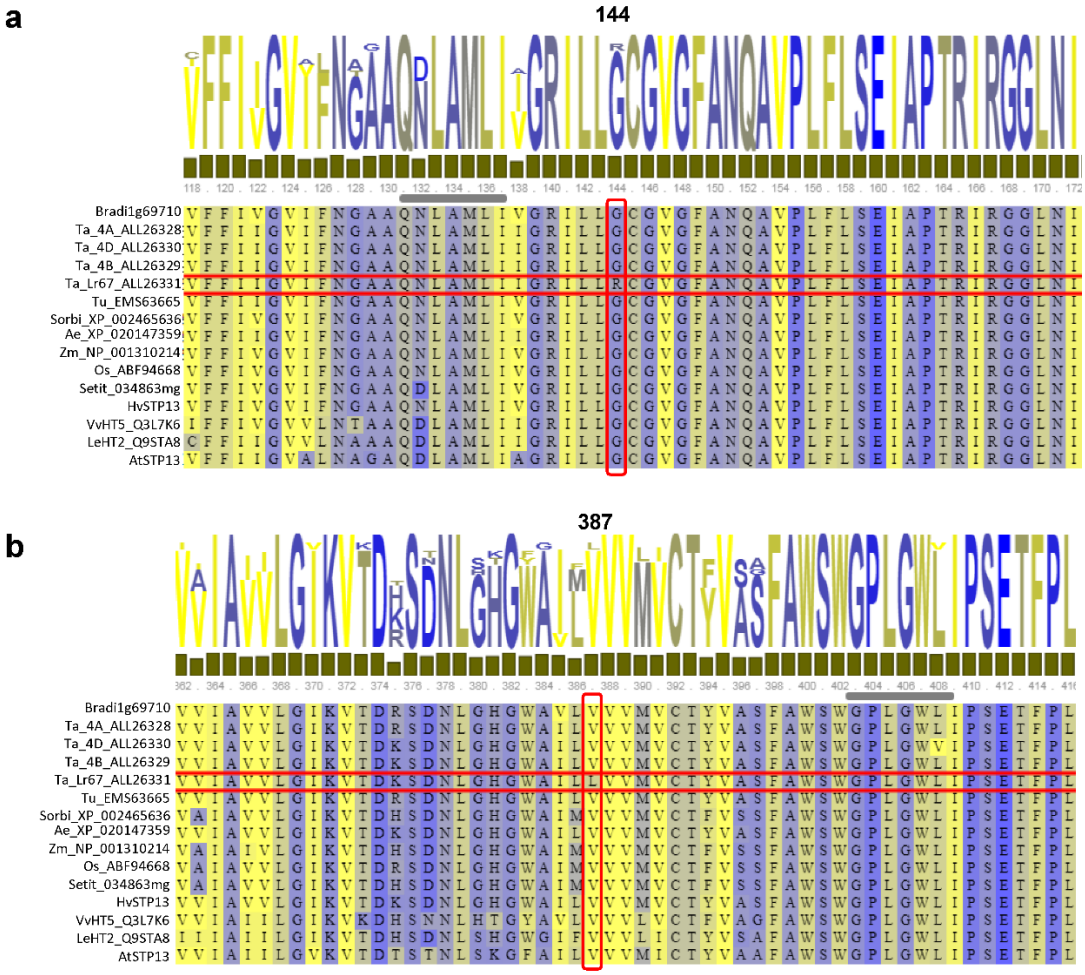


Figure 3.6. Protein alignment of STP13-like family members from different plant species including *B. distachyon*.
Species and proteins included for comparison are *B. distachyon* (Bradi1g69710, BdSTP13), *Triticum aestivum* (Ta_4A_ALL26328, Ta_4D_ALL26330 (*Lr67sus* allele), Ta_4B_ALL26329, Ta_Lr67_ALL26331 (*Lr67res* allele)), *Triticum urartu* (Tu_EMS63665), *Sorghum bicolor* (Sorbi_XP_002465636), *Aegilops tauschii* (Ae_XP_020147359), *Zea* mays (Zm_NP_001310214), *Oryza sativa* (Os_ABF94668), *Setaria italica* (Setit_034863mg), *Hordeum vulgare* (HvSTP13), *Vitis vinifera* (VvHT5_Q3L7K6), *Lycopersicon esculentum* (LeHT2_Q9STA8), *Arabidopsis thaliana* (AtSTP13). **a)** Close-up view of sequence alignment focusing on the amino acid sequence of BdSTP13 (Bradi1g69710) from position 118 to 172. Vertical red box highlights conservation of Gly144 across most proteins, while the horizontal red box shows protein sequence of Lr67res (Ta_Lr67_ALL26331) with a polymorphism (Arg, R) at position 144. **b)** Close-up view of sequence alignment zooms focusing on the amino acid sequence of BdSTP13 (Bradi1g69710) from position 362 to 412. Vertical red box highlights conservation of Val387 across most proteins while the horizontal red box shows protein sequence of Lr67res with a polymorphism (Leu, L) at position387.

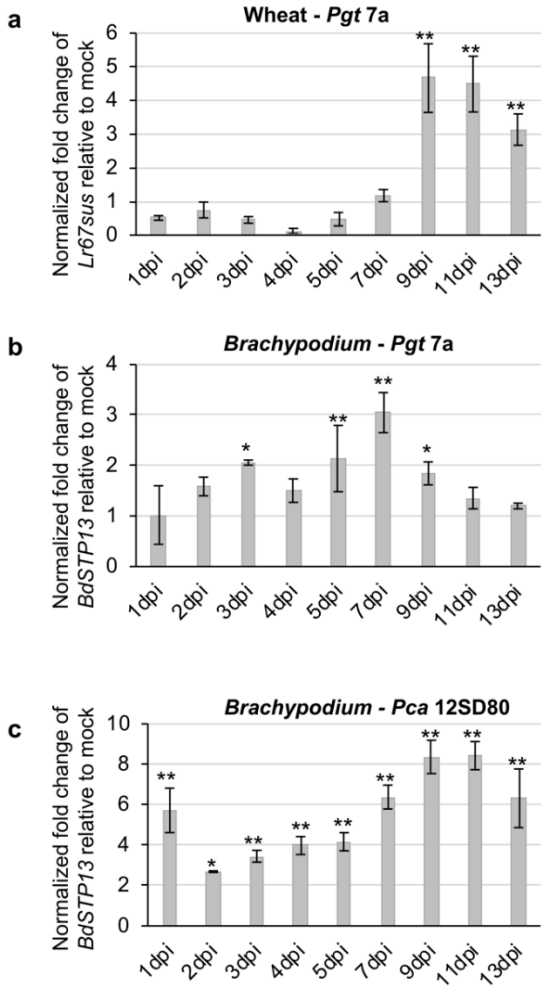


Figure 3.7. Time-course analysis of *BdSTP13* and *Lr67sus* gene expression upon infection with rust fungi.
a) Quantification of gene *Lr67sus* (NCBI accession: KR604816) expression in susceptible wheat cultivar McNair infected with *Pgt* isolate 7a. Gene expression was normalized using the wheat *GAPDH* (EU022331) as a reference gene. **b**) *BdSTP13* (*Bradi1g69710*) expression in WT Bd21-3 after infection with *Pgt* isolate 7a. **c)** *BdSTP13* expression in WT Bd21-3 after infection with *Pca* isolate 12SD80. For **b)** and **c)**, gene expression was normalized using *B. distachyon GAPDH* as a reference gene (*Bradi3g14120*). Y-axis represents fold change of gene expression values measured by RT-qPCR relative to the mean of mock-treated plants. Data represent mean ± standard error of the mean (SEM) of three biological replicates. Statistical analysis was performed using one-way ANOVA followed by multiple comparisons to define significance gene expression changes between infected plants and mock inoculation. Two asterisks represent p value < 0.01 while one asterisk represents p value < 0.05.

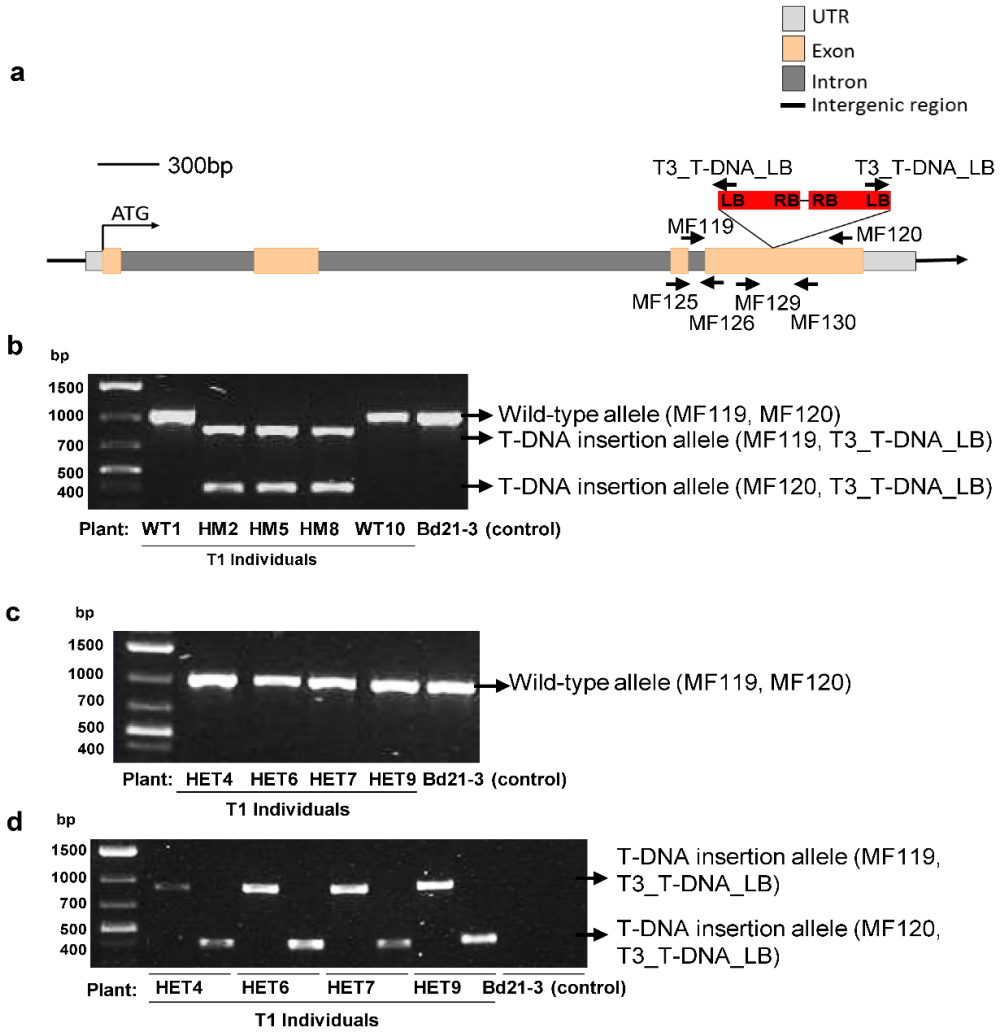


Figure 3.8. PCR-based T-DNA mutant genotyping of line JJ16783.
a) Diagram of insertion of T-DNA fragment in *Bradi1g69710* (*BdSTP13*) gene. The boxed region indicates the gene transcript with the untranslated region (UTR) in light grey box, coding sequence in orange and introns in dark grey. Backbone arrow depicts 5'-3' orientation. The location (Bd1: 68,416,145) of the start codon (ATG) is also shown above the gene model. The T-DNA insertion site (Bd1: 68,420,553) is illustrated by a red box above the fourth exon. Short arrows illustrate primers used for T-DNA genotyping and RT-PCR. **b)** to **d)** PCR-based genotyping analysis of T_1 individuals (line JJ16783) using gene-specific primers MF119, MF120 to target amplification of wild-type gene and the T-DNA primer T3_T-DNA_LB combined with primer MF119 and/or MF120 to target flanking regions of the insert. Gel electrophoresis images show PCR amplicons and GeneRuler 1kb DNA ladder (shown in the first lane). DNA template from wild type Bd21-3 was used as positive experimental control. Genotypes of T_1 individual plants are defined as WT to indicate homozygosity of the wild-type gene, HM to indicate homozygosity of disrupted gene; and HET to indicate heterozygosity of the wild-type gene as the disrupted gene is also present.

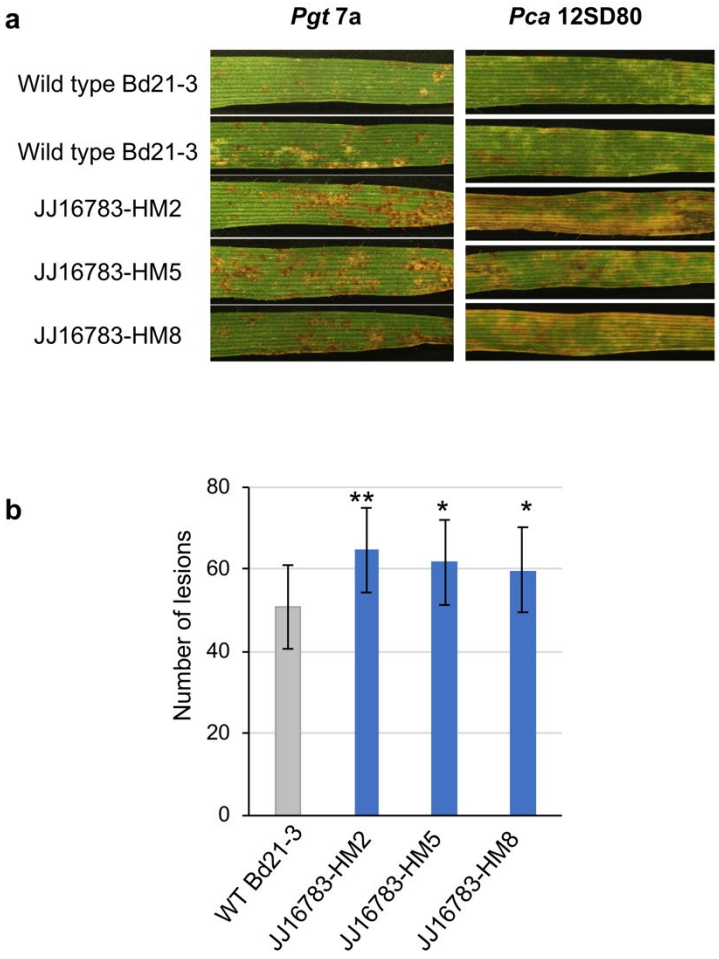


Figure 3.9. Phenotypic characterization of homozygous mutant individuals of *B. distachyon* line JJ16783 after challenge with rust pathogens.
a) Macroscopic symptoms in wild type Bd21-3 plants or T_2 homozygous mutant plants after infection with *Puccinia graminis* f. sp. *tritici* (*Pgt*) isolate 7a or *P. coronata* f. sp. *avenae* (*Pca*) isolate 12SD80. Photographs were taken at 14 dpi in both cases. **b)** Total number of macroscopic lesions in *Pgt* 7a-infected WT and T_2 homozygous mutants at 14 dpi. Data represent mean ± standard error of the mean (SEM) of two independent inoculation experiments. Each independent experiment consisted of at least 19 individual plants for each genotype (wild type Bd21-3, HM2, HM5, or HM8). Significant differences were evaluated using linear mixed-effects models (lme4) in R in comparison to WT Bd21-3. Two asterisks represent p value < 0.01 while one asterisk represents p value < 0.05.

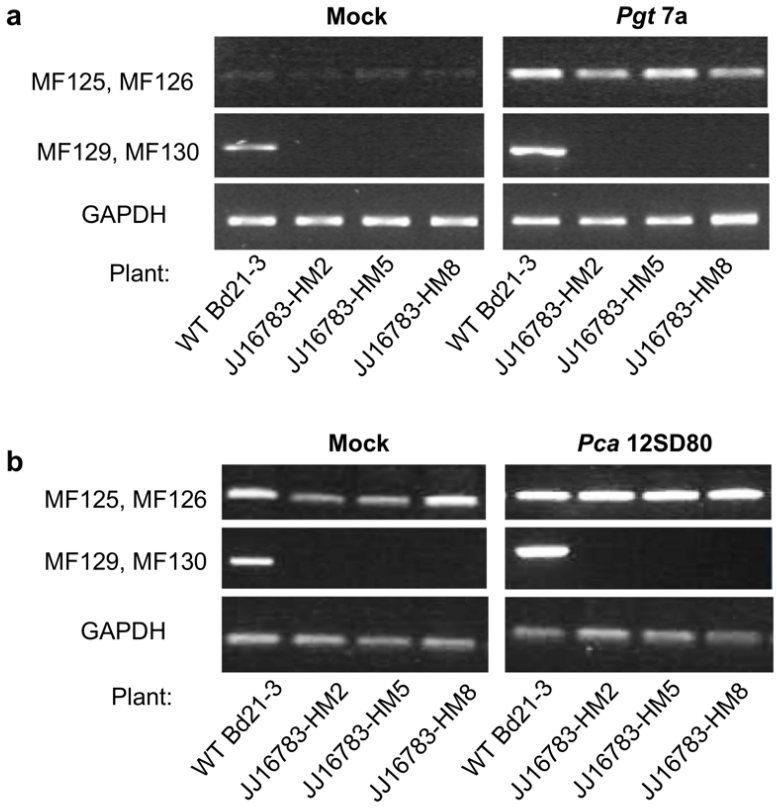


Figure 3.10. Reverse-transcription PCR (RT-PCR) analysis of *BdSTP13* gene expression in homozygous mutant lines of JJ16783 after infection with *P. graminis* f. sp. *tritici* (*Pgt*) or *P. coronata* f. sp. *avenae* (*Pca*).
a) *BdSTP13* expression in WT Bd21-3 or T_2 homozygous mutant plants (HM2, HM5, MH8) four days after *Pgt* 7a-infection or mock treatment. **b)** *BdSTP13* expression in WT Bd21-3 or T_2 homozygous mutant plants (HM2, HM5, MH8) four days after *Pca* 12SD80-infection or mock treatment. Primers MF125, MF126 were used to amplify the transcript region upstream of the T-DNA insertion site and primers MF129, MF130 were used to amplify region spanning the T-DNA insertion. Locations of primers were indicated in Figure 3.8a. Expression of *Brachypodium GAPDH* gene was used as positive experimental control.

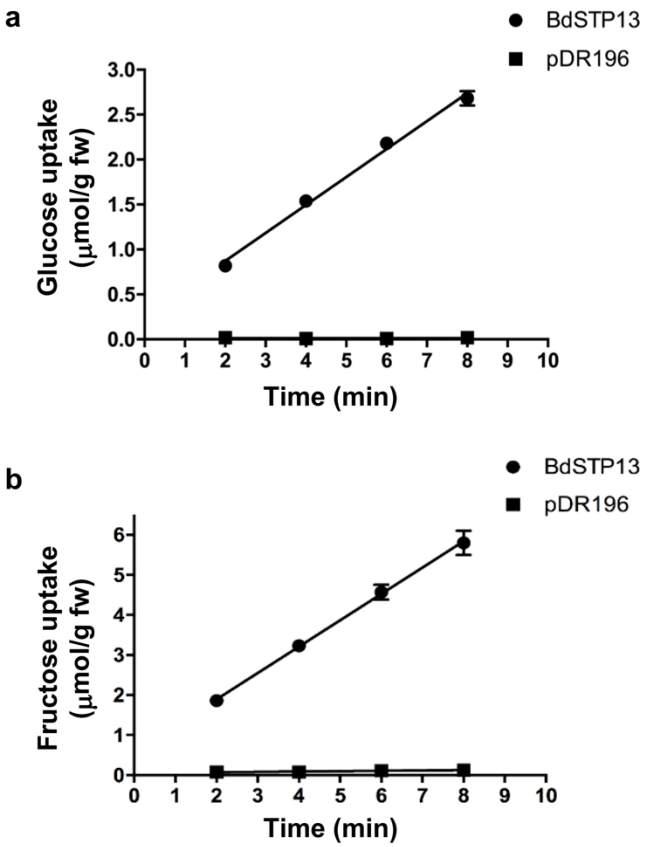


Figure 3.11. Time-course analysis of hexose uptake of heterologous expression of *BdSTP13* cDNA in yeast.

a) Line chart indicates glucose accumulation (µmol/grams of fresh weight (g fw) of yeast cells) in yeast cultures (strain EBY. VW4000) expressing *BdSTP13* (circle points) at 2 min time intervals after exposure to 100 µM of [^{14}C] glucose. Glucose accumulation values for yeast cultures containing the empty pDR196 vector (square points) are also shown (negative control). **b)** Line chart indicates fructose accumulation (µmol/g fw) in yeast cultures expressing *BdSTP13* (circle points) at 2 min time intervals after exposure to 1 mM of [^{14}C] fructose. Fructose accumulation values for yeast cultures containing the empty pDR196 vector (square points) are also shown (negative control). **a)** to **b)** Data represent mean ± SEM of three independent yeast cultures each containing two technical replicates.

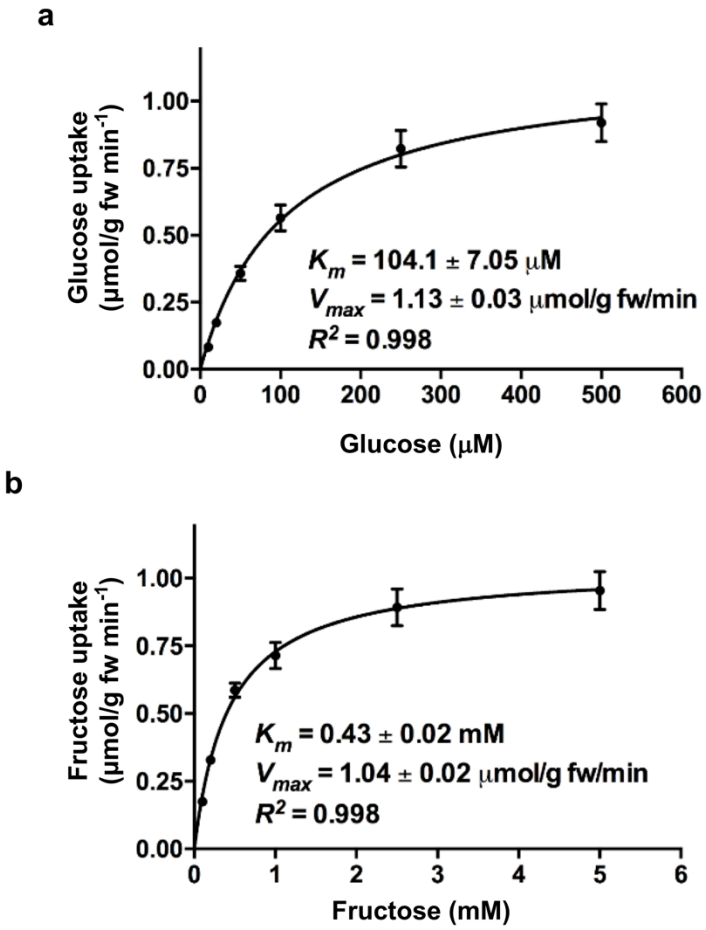


Figure 3.12. Transport affinity of hexose in yeast expressing *BdSTP13*.
a) Kinetic analysis of hexose uptake rate (µmol/g fw min^{-1}) of *BdSTP13* after incubation with various concentrations (µM) of [^{14}C] glucose. **b)** Kinetic analysis of hexose uptake rate (µmol/g fw min^{-1}) of *BdSTP13* after incubation with various concentrations (mM) of [^{14}C] fructose. **a)** and **b)** Data represent mean ± SEM of three independent yeast cultures each containing two technical replicates. Kinetics parameter V_{max} indicates the maximum rate of transport, while K_m shows the substrate concentration when reaching $V_{max}/2$. The degree of fit (R^2) is shown using Michaelis-Menten equation with least squares fit.

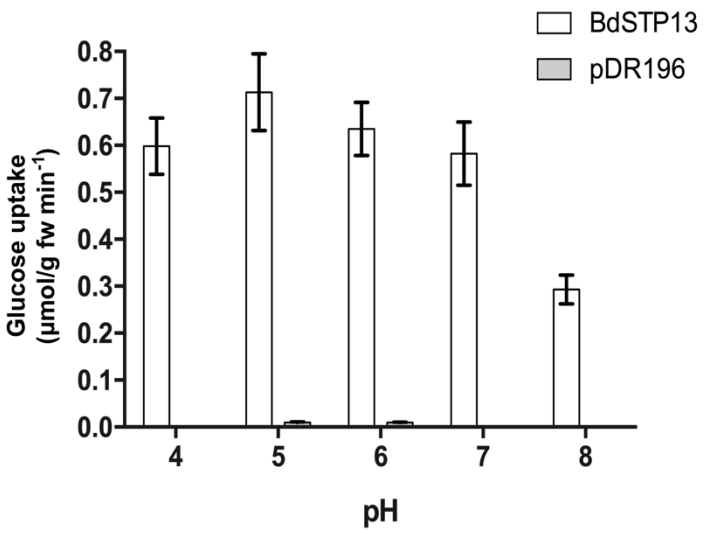


Figure 3.13. pH dependence of glucose transport rate of *BdSTP13* in yeast.
Bar graph shows glucose uptake rate (µmol/g fw min^{-1}) of *BdSTP13* in yeast cultures (white bars) and pDR196 empty vector (grey bars) pre-treated in 25 mM MES-HEPES solution at different pH levels. pDR196 backbone was included as a negative control. For all experiments, yeast cells were collected after incubation with 100 µM of [^{14}C] glucose solution. Data represent mean ± SEM of three independent yeast cultures, each containing two technical replicates.

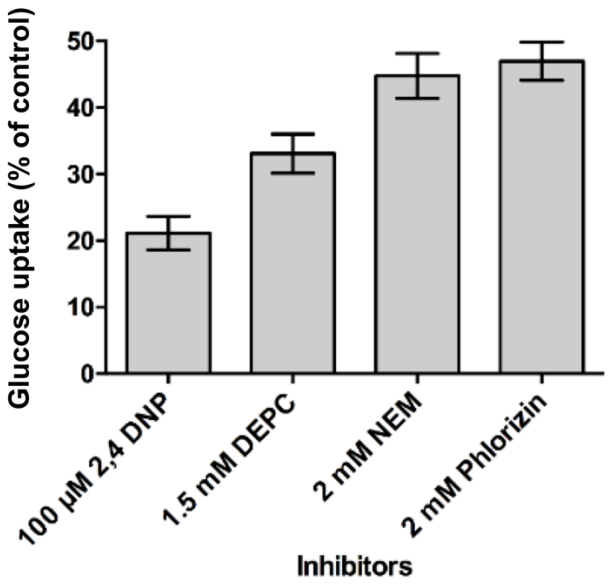


Figure 3.14. Effect of various inhibitors on glucose transport of *BdSTP13* in yeast. Bar graph shows the percentage of glucose uptake by *BdSTP13*-expressing yeast cells pre-incubated with 10 µl of inhibitors for 30 seconds (100 µM of 2,4-dinitrophenol (DNP), a proton ionophore; 1.5 mM of diethyl pyrocarbonate (DEPC) and 2 mM of N-ethylmaleimide (NEM), two sulfhydryl modifiers; 2 mM of phlorizin, a non-transported glucose analog inhibitor). Yeast culture without any addition of inhibitors served as experimental control and its glucose uptake was used as baseline to calculate percentage of uptake reduction. For all experiments, yeast cells were collected after incubation with 100 µM of [^{14}C] glucose solution (pH = 5). Data represent mean ± SEM of three independent yeast cultures, each containing two technical replicates.

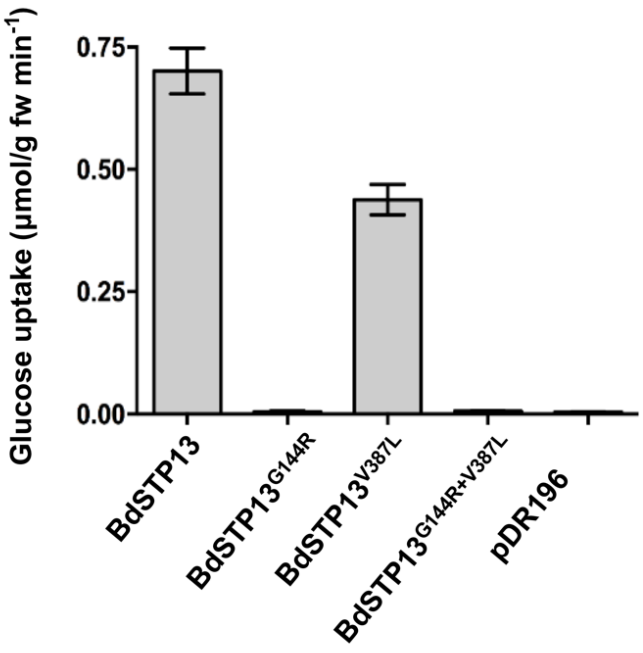


Figure 3.15. Glucose uptake rate of wild-type *BdSTP13* and site-directed mutants (G144R, V387L, G144R+V387L) in yeast.

Bar graph shows mean ± SEM of glucose uptake rate of yeast cells expressing *BdSTP13* or variants with single or double amino acid changes (*BdSTP13*^{G144R}, *BdSTP13*^{V387L} or *BdSTP13*$^{G144R+V387L}$). Yeast cells expressing the pDR196 vector backbone were included as a negative control. For all experiments, yeast cells were collected after incubation with 100 µM of [^{14}C] glucose solution (pH = 5). Data represent mean ± SEM of three independent assays, each containing two technical replicates.

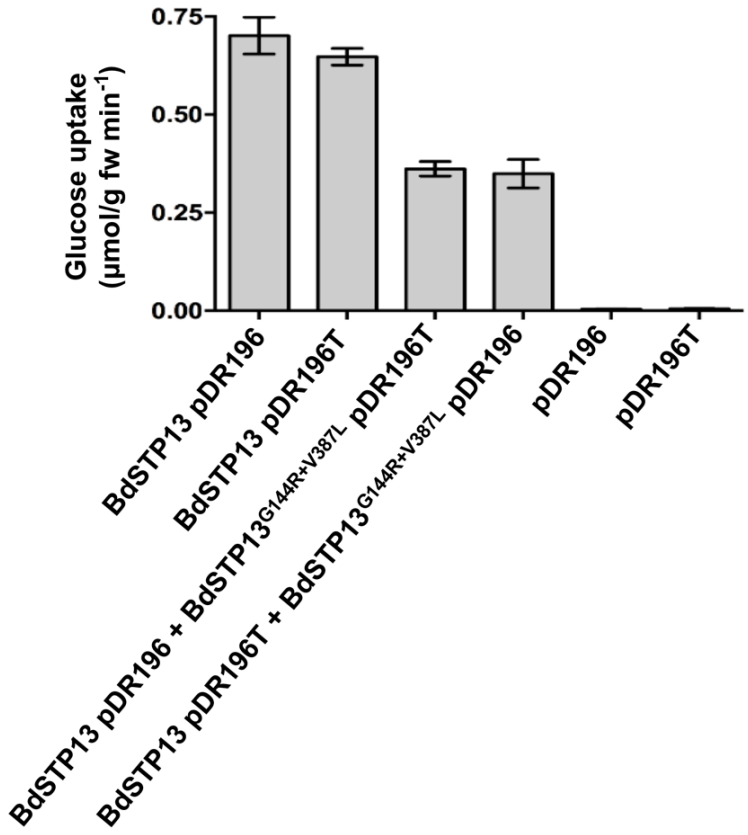


Figure 3.16. Effect of mutant *BdSTP13* (G144R+V387L) expression on uptake rate of wild-type transporter in yeast.

Bar graph shows mean ± SEM of glucose uptake rate of yeast cells expressing wild-type *BdSTP13* alone or co-expressing wild-type *BdSTP13* with the double mutant *BdSTP13*$^{G144R+V387L}$, or the empty vector controls pDR196 or pDR196T. Wild-type and mutant proteins were expressed from either the pDR196 or pDR196T vector backbones as indicated. For all experiments, yeast cells were collected after incubation with 100 µM of [^{14}C] glucose solution (pH = 5). Data represent mean ± SEM of three independent yeast cultures, each containing two technical replicates.

Bibliography

Agrios, G. N. (1997). *Plant pathology*. Academic press.

Aime, M. C., McTaggart, A. R., Mondo, S. J., and Duplessis, S. (2017). Phylogenetics and Phylogenomics of Rust Fungi. *Adv. Genet.* 100, 267–307. doi:10.1016/bs.adgen.2017.09.011.

Alexopoulos CJ., Mims CW., and Blackwell M. (1996). "Introductory Mycology. 4th ed.," in *Introductory Mycology. 4th ed.* (New York, USA: John Wiley & Sons), 868.

Anderson, C., Khan, M. A., Catanzariti, A.-M., Jack, C. A., Nemri, A., Lawrence, G. J., et al. (2016). Genome analysis and avirulence gene cloning using a high-density RADseq linkage map of the flax rust fungus, *Melampsora lini*. *BMC Genomics* 17, 667.

Anikster, Y. (1984). "4 - The Formae Speciales," in *The Cereal Rusts*, eds. W. R. Bushnell and A. P. Roelfs (Academic Press), 115–130. doi:10.1016/B978-0-12-148401-9.50010-8.

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., et al. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* 37, 139. doi:10.1038/s41587-018-0007-9.

Ayliffe, M., Devilla, R., Mago, R., White, R., Talbot, M., Pryor, A., et al. (2011). Nonhost resistance of rice to rust pathogens. *Mol. Plant. Microbe Interact.* 24, 1143–1155. Available at: http://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-04-11-0100 [Accessed December 11, 2014].

Ayliffe, M., Jin, Y., Kang, Z., Persson, M., Steffenson, B., Wang, S., et al. (2010). Determining the basis of nonhost resistance in rice to cereal rusts. *Euphytica* 179, 33–40. doi:10.1007/s10681-010-0280-2.

Ayliffe, M., Periyannan, S. K., Feechan, A., Dry, I., Schumann, U., Wang, M.-B., et al. (2013a). A Simple Method for Comparing Fungal Biomass in Infected Plant Tissues. *Mol. Plant. Microbe Interact.* 26, 658–667. doi:10.1094/MPMI-12-12-0291-R.

Ayliffe, M., Singh, D., Park, R., Moscou, M., and Pryor, T. (2013b). Infection of *Brachypodium distachyon* with Selected Grass Rust Pathogens. *Mol. Plant. Microbe Interact.* 26, 946–957. doi:10.1094/MPMI-01-13-0017-R.

Ayliffe, M., Singh, R., and Lagudah, E. (2008). Durable resistance to wheat stem rust needed. *Curr. Opin. Plant Biol.* 11, 187–192. doi:10.1016/j.pbi.2008.02.001.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi:10.1016/j.jml.2007.12.005.

Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* 6. doi:10.1186/s13100-015-0041-9.

Barbieri, M., Marcel, T. C., and Niks, R. E. (2011). Host Status of False Brome Grass to the Leaf Rust Fungus *Puccinia brachypodii* and the Stripe Rust Fungus *P. striiformis*. *Plant Dis.* 95, 1339–1345. doi:10.1094/PDIS-11-10-0825.

Barrett, L. G., Thrall, P. H., Dodds, P. N., van der Merwe, M., Linde, C. C., Lawrence, G. J., et al. (2009). Diversity and Evolution of Effector Loci in Natural Populations of the Plant Pathogen *Melampsora lini*. *Mol. Biol. Evol.* 26, 2499–2513. doi:10.1093/molbev/msp166.

Bartos, P., Fleischmann, G., Samborski, D., and Shipton, W. (1969). Studies on asexual variation in the virulence of oat crown rust, *Puccinia coronata* f. sp. *avenae*, and wheat leaf rust, *Puccinia recondita*. *Can. J. Bot.* 47, 1383–1387.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version* 1, 1–23.

Belton, J. M., McCord, R. P., Gibcus, J. H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* 58, 268–76. doi:10.1016/j.ymeth.2012.05.001.

Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340, 783–795.

Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., et al. (2011). Structural and Functional Analysis of a Plant Resistance Protein TIR Domain Reveals Interfaces for Self-Association, Signaling, and Autoregulation. *Cell Host Microbe* 9, 200–211. doi:10.1016/j.chom.2011.02.009.

Bettgenhaeuser, J., Gilbert, B., Ayliffe, M., and Moscou, M. J. (2014). Nonhost resistance to rust pathogens – a continuation of continua. *Front. Plant Sci.* 5. doi:10.3389/fpls.2014.00664.

Bhattacharya, S. (2017). Deadly new wheat disease threatens Europe's crops. *Nat. News* 542, 145. doi:10.1038/nature.2017.21424.

Birkenbihl, R. P., Diezel, C., and Somssich, I. E. (2012). Arabidopsis WRKY33 Is a Key Transcriptional Regulator of Hormonal and Metabolic Responses toward *Botrytis cinerea* Infection. *Plant Physiol.* 159, 266–285. doi:10.1104/pp.111.192641.

Boehm, E. W. A., Wenstrom, J. C., McLaughlin, D. J., Szabo, L. J., Roelfs, A. P., and Bushnell, W. R. (1992). An ultrastructural pachytene karyotype for *Puccinia graminis* f. sp. *tritici*. *Can. J. Bot.* 70, 401-413.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.

Bolton, M. D., Kolmer, J. A., and Garvin, D. F. (2008). Wheat leaf rust caused by *Puccinia triticina*. *Mol. Plant Pathol.* 9, 563–575. doi:10.1111/j.1364-3703.2008.00487.x.

Borlaug, N. E. (1968). Wheat breeding and its impact on world food supply. Australian Academy of Science Canberra Available at: http://libcatalog.cimmyt.org/download/borlaug/66179.pdf [Accessed October 22, 2016].

Bragg, J. N., Wu, J., Gordon, S. P., Guttman, M. E., Thilmony, R., Lazo, G. R., et al. (2012). Generation and Characterization of the Western Regional Research Center Brachypodium T-DNA Insertional Mutant Collection. *PLoS ONE* 7, e41916. doi:10.1371/journal.pone.0041916.

Brkljacic, J., Grotewold, E., Scholl, R., Mockler, T., Garvin, D. F., Vain, P., et al. (2011). Brachypodium as a Model for the Grasses: Today and the Future. *PLANT Physiol.* 157, 3–13. doi:10.1104/pp.111.179531.

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi:10.1038/nmeth.3176.

Burdon, J. J., Marshall, D. R., and Luig, N. H. (1981). Isozyme analysis indicates that a virulent cereal rust pathogen is a somatic hybrid. *Nature* 293, 565. doi:10.1038/293565a0.

Burdon, J. J., Marshall, D. R., Luig, N. H., and Gow, D. J. S. (1982). Isozyme Studies on the Origin and Evolution of *Puccinia graminis* f. sp. *tritici* in Australia. *Aust. J. Biol. Sci.* 35, 231–238. Available at: https://doi.org/10.1071/BI9820231.

Burdon, J. J., and Silk, J. (1997). Sources and Patterns of Diversity in Plant-Pathogenic Fungi. *Phytopathology* 87, 664–669. doi:10.1094/PHYTO.1997.87.7.664.

Burrage, S. W. (1970). Environmental factors influencing the infection of wheat by *Puccinia graminis*. *Ann. Appl. Biol.* 66, 429–440. doi:10.1111/j.1744-7348.1970.tb04622.x.

Bush, D. R. (1993). Proton-Coupled Sugar and Amino Acid Transporters in Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44, 513–542. doi:10.1146/annurev.pp.44.060193.002501.

Büttner, M. (2010). The Arabidopsis sugar transporter (AtSTP) family: an update. *Plant Biol.* 12, 35–41. doi:10.1111/j.1438-8677.2010.00383.x.

Cabanettes, F., and Klopp, C. (2018). D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ* 6, e4958. doi:10.7717/peerj.4958.

Casey, L. W., Lavrencic, P., Bentham, A. R., Cesari, S., Ericsson, D. J., Croll, T., et al. (2016). The CC domain structure from the wheat stem rust resistance protein *Sr33* challenges paradigms for dimerization in plant NLR proteins. *Proc. Natl. Acad. Sci.* 113, 12856–12861. doi:10.1073/pnas.1609922113.

Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: the 'integrated decoy' hypothesis. *Front. Plant Sci.* 5. doi:10.3389/fpls.2014.00606.

Chakraborty, S., Luck, J., Hollaway, G., Fitzgerald, G., and White, N. (2011). Rust-proofing wheat for a changing climate. *Euphytica* 179, 19–32. doi:10.1007/s10681-010-0324-7.

Chen, J., Upadhyaya, N. M., Ortiz, D., Sperschneider, J., Li, F., Bouton, C., et al. (2017). Loss of *AvrSr50* by somatic exchange in stem rust leads to virulence for *Sr50* resistance in wheat. *Science* 358, 1607–1610. doi:10.1126/science.aao4810.

Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., et al. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468, 527–532. doi:10.1038/nature09606.

Chen, S., Rouse, M. N., Zhang, W., Zhang, X., Guo, Y., Briggs, J., et al. (2019). Wheat gene *Sr60* encodes a protein with two putative kinase domains that confers resistance to stem rust. *New Phytol.* doi:10.1111/nph.16169.

Chen, S., Zhang, W., Bolus, S., Rouse, M. N., and Dubcovsky, J. (2018). Identification and characterization of wheat stem rust resistance gene *Sr21* effective against the Ug99 race group at high temperature. *PLOS Genet.* 14, e1007287. doi:10.1371/journal.pgen.1007287.

Chen, W. Q., Wellings, C., Chen, X. M., Kang, Z. S., and Liu, T. G. (2014). Wheat stripe (yellow) rust caused by *Puccinia striiformis* f. sp. *tritici*. *Mol Plant Pathol* 15, 433–446. doi:10.1111/mpp.12116.

Chen, X. M. (2005). Epidemiology and control of stripe rust *Puccinia striiformis* f. sp. *tritici* on wheat. *Can. J. Plant Pathol.* 27, 314–337. doi:10.1080/07060660509507230.

Cheng, Y., Zhang, H., Yao, J., Wang, X., Xu, J., Han, Q., et al. (2012). Characterization of non-host resistance in broad bean to the wheat stripe rust pathogen. *BMC Plant Biol.* 12, 96. doi:10.1186/1471-2229-12-96.

Chojnacki, S., Cowley, A., Lee, J., Foix, A., and Lopez, R. (2017). Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. *Nucleic Acids Res.* 45, W550–W553. doi:10.1093/nar/gkx273.

CIMMYT (2017). WHEAT in the World. Available at: https://wheat.org/wheat-in-the-world/.

Clark, T. A., and Anderson, J. B. (2004). Dikaryons of the Basidiomycete Fungus *Schizophyllum commune*: Evolution in Long-Term Culture. *Genetics* 167, 1663–75. doi:10.1534/genetics.104.027235.

Collins, N. C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J.-L., et al. (2003). SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* 425, 973. doi:10.1038/nature02076.

Consortium, U. (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 46, 2699.

Corradi, N., and Brachmann, A. (2017). Fungal Mating in the Most Widespread Plant Symbionts? *Trends Plant Sci.* 22, 175–183. doi:10.1016/j.tplants.2016.10.010.

Crowley, P. H. (1975). Natural selection and the Michaelis constant. *J. Theor. Biol.* 50, 461–475. doi:10.1016/0022-5193(75)90093-4.

Cuomo, C. A., Bakkeren, G., Khalil, H. B., Panwar, V., Joly, D., Linning, R., et al. (2017). Comparative Analysis Highlights Variable Genome Content of Wheat Rusts and Divergence of the Mating Loci. *G3-Genes Genomes Genet.* 7, 361–376. doi:10.1534/g3.116.032797.

Davies K., Evans A., and Oxley S. (2007). Impact Of Climate Change In Scotland On Crop Pests, Weeds And Disease - SRUC. Available at: https://www.sruc.ac.uk/downloads/file/766/tn605_impact_of_climate_change_in_scotland_on_crop_pests_weeds_and_disease [Accessed March 3, 2019].

Dawson, A. M., Bettgenhaeuser, J., Gardiner, M., Green, P., Hernández-Pinzón, I., Hubbard, A., et al. (2015). The development of quick, robust, quantitative phenotypic assays for describing the host–nonhost landscape to stripe rust. *Front. Plant Sci.* 6. doi:10.3389/fpls.2015.00876.

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., et al. (2012). The Top 10 fungal pathogens in molecular plant

pathology: Top 10 fungal pathogens. *Mol. Plant Pathol.* 13, 414–430. doi:10.1111/j.1364-3703.2011.00783.x.

DeYoung, B. J., and Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat. Immunol.* 7, 1243–1249. doi:10.1038/ni1410.

Dodds, P. N., and Lagudah, E. S. (2016). Starving the enemy. *Science* 354, 1377–1378. doi:10.1126/science.aal4273.

Dodds, P. N., Lawrence, G. J., Catanzariti, A.-M., Teh, T., Wang, C.-I., Ayliffe, M. A., et al. (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc. Natl. Acad. Sci.* 103, 8888–8893. Available at: http://www.pnas.org/content/103/23/8888.short [Accessed October 22, 2016].

Dodds, P. N., and Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nat. Rev. Genet.* 11, 539–548. doi:10.1038/nrg2812.

Dohmen, R. J., Strasser, A. W., Höner, C. B., and Hollenberg, C. P. (1991). An efficient transformation procedure enabling long-term storage of competent cells of various yeast genera. *Yeast* 7, 691–692.

Draper, J., Mur, L. A. J., Jenkins, G., Ghosh-Biswas, G. C., Bablak, P., Hasterok, R., et al. (2001). *Brachypodium distachyon*. A New Model System for Functional Genomics in Grasses. *PLANT Physiol.* 127, 1539–1555. doi:10.1104/pp.010196.

Duplessis, S., Bakkeren, G., and Hamelin, R. (2014). "Advancing knowledge on biology of rust fungi through genomics," in *Advances in Botanical Research* Fungi., ed. F. M. Martin (Academic Press), 173–209. doi:10.1016/B978-0-12-397940-7.00006-9.

Duplessis, S., Cuomo, C. A., Lin, Y.-C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., et al. (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9166–9171. doi:10.1073/pnas.1019315108.

Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19, 1349. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC333874/ [Accessed August 26, 2019].

Ellingboe, A. (1961). Somatic recombination in *Puccinia graminis* var. *tritici*. *Phytopathology* 51, 13–15.

Ellis, J. (2006). Insights into Nonhost Disease Resistance: Can They Assist Disease Control in Agriculture? *Plant Cell* 18, 523–528. doi:10.1105/tpc.105.040584.

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., and Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. *Front. Plant Sci.* 5. doi:10.3389/fpls.2014.00641.

Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16, 157. doi:10.1186/s13059-015-0721-2.

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5, 199–206. doi:10.1016/S1360-1385(00)01600-9.

Eulgem, T., Rushton, P. J., Schmelzer, E., Hahlbrock, K., and Somssich, I. E. (1999). Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *EMBO J.* 18, 4689–4699. doi:10.1093/emboj/18.17.4689.

Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10, 366–371. doi:10.1016/j.pbi.2007.04.020.

Eversmeyer, M. G., and Kramer, C. L. (2000). Epidemiology of Wheat Leaf and Stem Rust in the Central Great Plains of the USA. *Annu. Rev. Phytopathol.* 38, 491–513. doi:10.1146/annurev.phyto.38.1.491.

Fan, J., and Doerner, P. (2012). Genetic and molecular basis of nonhost disease resistance: complex, yes; silver bullet, no. *Curr. Opin. Plant Biol.* 15, 400–406. doi:10.1016/j.pbi.2012.03.001.

FAO (2017). FAOSTAT statistics database, Crops. Available at: www.fao.org/faostat/en/#data/QC.

Felle, H., Gogarten, J. P., and Bentrup, F. W. (1983). Phlorizin inhibits hexose transport across the plasmalemma of Riccia fluitans. *Planta* 157, 267–270. doi:10.1007/BF00405192.

Fetch, T., Zegeye, T., Park, R. F., Hodson, D., and Wanyera, R. (2016). Detection of Wheat Stem Rust Races TTHSK and PTKTK in the Ug99 Race Group in Kenya in 2014. *Plant Dis.* 100, 1495–1495. doi:10.1094/PDIS-11-15-1356-PDN.

Figueroa, M., Alderman, S., Garvin, D. F., and Pfender, W. F. (2013). Infection of *Brachypodium distachyon* by Formae Speciales of *Puccinia graminis*: Early Infection Events and Host-Pathogen Incompatibility. *PLoS ONE* 8, e56857. doi:10.1371/journal.pone.0056857.

Figueroa, M., Castell-Miller, C. V., Li, F., Hulbert, S. H., and Bradeen, J. M. (2015). Pushing the boundaries of resistance: insights from Brachypodium-rust interactions. *Plant Biot. Interact.*, 558. doi:10.3389/fpls.2015.00558.

Figueroa, M., Hammond-Kosack, K. E., and Solomon, P. S. (2018). A review of wheat diseases-a field perspective. *Mol Plant Pathol* 19, 1523–1536. doi:10.1111/mpp.12618.

Figueroa, M., Upadhyaya, N. M., Sperschneider, J., Park, R. F., Szabo, L. J., Steffenson, B., et al. (2016). Changing the Game: Using Integrative Genomics to Probe Virulence Mechanisms of the Stem Rust Pathogen *Puccinia graminis* f. sp. *tritici*. *Front. Plant Sci.* 7. doi:10.3389/fpls.2016.00205.

Fleißner, A., and Serrano, A. (2016). "The Art of Networking: Vegetative Hyphal Fusion in Filamentous Ascomycete Fungi," in *Growth, Differentiation and Sexuality* The Mycota., ed. J. Wendland (Cham: Springer International Publishing), 133–153. doi:10.1007/978-3-319-25844-7_7.

Flor, H. (1964). Genetics of somatic variation for pathogenicity in *Melampsora lini*. *Phytopathology* 54, 823–826.

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* 9, 275–296. Available at: http://www.annualreviews.org/doi/pdf/10.1146/annurev.py.09.090171.001423 [Accessed October 22, 2016].

Freeman, B. C., and Beattie, G. A. (2008). An overview of plant defenses against pathogens and herbivores. *Plant Health Instr.*

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., et al. (2009). A Kinase-START Gene Confers Temperature-Dependent Resistance to Wheat Stripe Rust. *Science* 323, 1357–1360. doi:10.1126/science.1166289.

Garnica, D. P., Nemri, A., Upadhyaya, N. M., Rathjen, J. P., and Dodds, P. N. (2014). The Ins and Outs of Rust Haustoria. *PLoS Pathog.* 10, e1004329. Available at: http://dx.plos.org/10.1371/journal.ppat.1004329 [Accessed December 11, 2014].

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *ArXiv12073907 Q-Bio*. Available at: http://arxiv.org/abs/1207.3907 [Accessed March 15, 2019].

Gates, J. E., and Loegering, W. Q. (1991). Mutation to wider virulence in *Puccinia graminis* f. sp. *tritici*: evidence for the existence of loci which allow the fungus to overcome several host stem rust resistance genes simultaneously. *Appl Env. Microbiol* 57, 2332–2336.

Gear, M. L., McPhillips, M. L., Patrick, J. W., and McCurdy, D. W. (2000). Hexose transporters of tomato: molecular cloning, expression analysis and functional characterization. *Plant Mol. Biol.* 44, 687–697.

Gel, B., and Serra, E. (2017). karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 33, 3088–3090.

German, S. E., and Kolmer, J. A. (1992). Effect of gene *Lr34* in the enhancement of resistance to leaf rust of wheat. *TAG Theor. Appl. Genet. Theor. Angew. Genet.* 84, 97–105. doi:10.1007/BF00223987.

Ghurye, J., Pop, M., Koren, S., Bickhart, D., and Chin, C.-S. (2017). Scaffolding of long read assemblies using long range contact information. *BMC Genomics* 18, 527.

Glazebrook, J. (2005). Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annu. Rev. Phytopathol.* 43, 205–227. doi:10.1146/annurev.phyto.43.040204.135923.

González-Lamothe, R., Mitchell, G., Gattuso, M., Diarra, M., Malouin, F., and Bouarab, K. (2009). Plant antimicrobial agents and their effects on plant and human pathogens. *Int. J. Mol. Sci.* 10, 3400–3419.

Gordon, S. P., Contreras-Moreira, B., Woods, D. P., Marais, D. L. D., Burgess, D., Shu, S. Q., et al. (2017). Extensive gene content variation in the *Brachypodium distachyon* pan-genome correlates with population structure. *Nat Commun* 8. doi:Artn 2184 10.1038/S41467-017-02292-8.

Gordon, S. P., Priest, H., Des Marais, D. L., Schackwitz, W., Figueroa, M., Martin, J., et al. (2014). Genome diversity in *Brachypodium distachyon:* deep sequencing of highly diverse inbred lines. *Plant J.* 79, 361–374. doi:10.1111/tpj.12569.

Gou, J.-Y., Li, K., Wu, K., Wang, X., Lin, H., Cantu, D., et al. (2015). Wheat Stripe Rust Resistance Protein WKS1 Reduces the Ability of the Thylakoid-Associated Ascorbate Peroxidase to Detoxify Reactive Oxygen Species. *Plant Cell* 27, 1755–1770. doi:10.1105/tpc.114.134296.

Govrin, E. M., and Levine, A. (2000). The hypersensitive response facilitates plant infection by the necrotrophic pathogen *Botrytis cinerea*. *Curr. Biol. CB* 10, 751–757. doi:10.1016/s0960-9822(00)00560-1.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi:10.1038/nbt.1883.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi:10.1093/bioinformatics/btt086.

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., et al. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* 9, R7. doi:10.1186/gb-2008-9-1-r7.

Hart, H. (1929). Relation of stomatal behavior to stem rust resistance in wheat. *J Agric Res* 39, 929–48.

Hayes, M. A., Feechan, A., and Dry, I. B. (2010). Involvement of abscisic acid in the coordinated regulation of a stress-inducible hexose transporter (VvHT5) and a cell wall invertase in grapevine in response to biotrophic fungal infection. *Plant Physiol.* 153, 211–221.

Heath, M. C. (2000). Nonhost resistance and nonspecific plant defenses. *Curr. Opin. Plant Biol.* 3, 315–319.

Herbers, K., Meuwly, P., Frommer, W. B., Metraux, J. P., and Sonnewald, U. (1996). Systemic Acquired Resistance Mediated by the Ectopic Expression of Invertase: Possible Hexose Sensing in the Secretory Pathway. *Plant Cell* 8, 793–803. doi:10.1105/tpc.8.5.793.

Hodson, D. P. (2011). Shifting boundaries: challenges for rust monitoring. *Euphytica* 179, 93–104. doi:10.1007/s10681-010-0335-4.

Hsia, M. M., O'Malley, R., Cartwright, A., Nieu, R., Gordon, S. P., Kelly, S., et al. (2017). Sequencing and functional validation of the JGI *Brachypodium distachyon* T-DNA collection. *Plant J.* 91, 361–370. doi:10.1111/tpj.13582.

Hsing, Y.-I., Chern, C.-G., Fan, M.-J., Lu, P.-C., Chen, K.-T., Lo, S.-F., et al. (2007). A rice gene activation/knockout mutant resource for high throughput functional genomics. *Plant Mol. Biol.* 63, 351–364. doi:10.1007/s11103-006-9093-z.

Hückelhoven, R. (2007). Cell wall–associated mechanisms of disease resistance and susceptibility. *Annu Rev Phytopathol* 45, 101–127. doi:10.1146/annurev.phyto.45.062806.094325.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2015). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi:10.1093/nar/gkv1248.

International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463, 763. doi:10.1038/nature08747.

Jin, Y. (2011). Role of Berberis spp. as alternate hosts in generating new races of *Puccinia graminis* and *P. striiformis*. *Euphytica* 179, 105–108. doi:10.1007/s10681-010-0328-3.

Jin, Y., Szabo, L. J., Pretorius, Z. A., Singh, R. P., Ward, R., and Fetch, T. (2008). Detection of Virulence to Resistance Gene *Sr24* Within Race TTKS of *Puccinia graminis* f. sp. *tritici*. *Plant Dis.* 92, 923–926. doi:10.1094/PDIS-92-6-0923.

Johnson, R. (1984). A Critical Analysis of Durable Resistance. *Annu. Rev. Phytopathol.* 22, 309–330. doi:10.1146/annurev.py.22.090184.001521.

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi:10.1038/nature05286.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinforma. Oxf. Engl.* 30, 1236–1240. doi:10.1093/bioinformatics/btu031.

Katsuya, K., and Green, G. J. (1967). Reproductive Potentials of Races 15b and 56 of Wheat Stem Rust. *Can. J. Bot.* 45, 1077-. doi:DOI 10.1139/b67-113.

Kawashima, C. G., Guimaraes, G. A., Nogueira, S. R., MacLean, D., Cook, D. R., Steuernagel, B., et al. (2016). A pigeonpea gene confers resistance to Asian soybean rust in soybean. *Nat Biotechnol* 34, 661–665. doi:10.1038/nbt.3554.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317.

Knaus, B. J., and Grünwald, N. J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53. doi:10.1111/1755-0998.12549.

Knott, D. R. (1989). "The Wheat Rust Pathogens," in *The Wheat Rusts — Breeding for Resistance* (Berlin, Heidelberg: Springer Berlin Heidelberg), 33–37. Available at: https://doi.org/10.1007/978-3-642-83641-1_2.

Kolmer, J. A. (2005). Tracking wheat rust on a continental scale. *Curr. Opin. Plant Biol.* 8, 441–449. doi:10.1016/j.pbi.2005.05.001.

Kolmer, J. A., and Peterson, P. M. (2001). "Early research on the genetics of *Puccinia graminis* and stem rust resistance in wheat in Canada and the United States," in *Stem rust of wheat: from ancient enemy to modern foe.* (APS Press), 53–60.

Available at: https://www.cabdirect.org/cabdirect/abstract/20013088372 [Accessed July 26, 2019].

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi:10.1101/gr.215087.116.

Krattinger, S. G., Jordan, D. R., Mace, E. S., Raghavan, C., Luo, M.-C., Keller, B., et al. (2013). Recent emergence of the wheat Lr34 multi-pathogen resistance: insights from haplotype analysis in wheat, rice, sorghum and *Aegilops tauschii*. *Theor. Appl. Genet.* 126, 663–672. doi:10.1007/s00122-012-2009-1.

Krattinger, S. G., Kang, J., Bräunlich, S., Boni, R., Chauhan, H., Selter, L. L., et al. (2019). Abscisic acid is a substrate of the ABC transporter encoded by the durable wheat disease resistance gene *Lr34*. *New Phytol.* doi:10.1111/nph.15815.

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., et al. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323, 1360–1363. doi:10.1126/science.1166453.

Krattinger, S. G., Lagudah, E. S., Wicker, T., Risk, J. M., Ashton, A. R., Selter, L. L., et al. (2011). Lr34 multi-pathogen resistance ABC transporter: molecular analysis of homoeologous and orthologous genes in hexaploid wheat and other grass species. *Plant J.* 65, 392–403. doi:10.1111/j.1365-313X.2010.04430.x.

Krattinger, S. G., Sucher, J., Selter, L. L., Chauhan, H., Zhou, B., Tang, M., et al. (2016). The wheat durable, multipathogen resistance gene Lr34 confers partial blast resistance in rice. *Plant Biotechnol. J.* 14, 1261–1268. doi:10.1111/pbi.12491.

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi:10.1006/jmbi.2000.4315.

Kruzel, E. K., and Hull, C. M. (2010). Establishing an unusual cell type: how to make a dikaryon. *Curr. Opin. Microbiol.* 13, 706–711.

Kurkdjian, A., and Guern, J. (1989). Intracellular pH: Measurement and Importance in Cell Activity. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 40, 271–303. doi:10.1146/annurev.pp.40.060189.001415.

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5. doi:Artn R12 Doi 10.1186/Gb-2004-5-2-R12.

Lacombe, S., Rougon-Cardoso, A., Sherwood, E., Peeters, N., Dahlbeck, D., van Esse, H. P., et al. (2010). Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. *Nat. Biotechnol.* 28, 365–369. doi:10.1038/nbt.1613.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., et al. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Comput. Biol.* 9, e1003118. doi:10.1371/journal.pcbi.1003118.

Lei, Y., Wang, M., Wan, A., Xia, C., See, D. R., Zhang, M., et al. (2017). Virulence and Molecular Characterization of Experimental Isolates of the Stripe Rust Pathogen (*Puccinia striiformis*) Indicate Somatic Recombination. *Phytopathology* 107, 329–344. doi:10.1094/PHYTO-07-16-0261-R.

Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., et al. (2019). NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res.* 47, W260–W265. doi:10.1093/nar/gkz303.

Lemoine, R., and Delrot, S. (1987). Recognition of phlorizin by the carriers of sucrose and hexose in broad bean leaves. *Physiol. Plant.* 69, 639–644. doi:10.1111/j.1399-3054.1987.tb01978.x.

Lemonnier, P., Gaillard, C., Veillet, F., Verbeke, J., Lemoine, R., Coutos-Thévenot, P., et al. (2014). Expression of Arabidopsis sugar transport protein STP13 differentially affects glucose transport activity and basal resistance to *Botrytis cinerea*. *Plant Mol. Biol.* 85, 473–484. doi:10.1007/s11103-014-0198-5.

Leonard, K. J., and Szabo, L. J. (2005). Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol. Plant Pathol.* 6, 99–111. doi:10.1111/j.1364-3703.2005.00273.x.

Lewis, C. M., Persoons, A., Bebber, D. P., Kigathi, R. N., Maintz, J., Findlay, K., et al. (2018). Potential for re-emergence of wheat stem rust in the United Kingdom. *Commun. Biol.* 1. doi:10.1038/s42003-018-0013-y.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinforma. Oxf. Engl.* 34, 3094–3100. doi:10.1093/bioinformatics/bty191.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics* 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

Li, J., Brader, G., Kariola, T., and Palva, E. T. (2006). WRKY70 modulates the selection of signaling pathways in plant defense. *Plant J. Cell Mol. Biol.* 46, 477–491. doi:10.1111/j.1365-313X.2006.02712.x.

Lipka, V., Dittgen, J., Bednarek, P., Bhat, R., Wiermer, M., Stein, M., et al. (2005). Pre- and Postinvasion Defenses Both Contribute to Nonhost Resistance in Arabidopsis. *Science* 310, 1180–1183. doi:10.1126/science.1119409.

Little, R., and Manners, J. G. (1969). Somatic recombination in yellow rust of wheat (*Puccinia striiformis*): I. The production and possible origin of two new physiologic races. *Trans. Br. Mycol. Soc.* 53, 251–258. doi:10.1016/S0007-1536(69)80059-8.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods San Diego Calif* 25, 402–408. doi:10.1006/meth.2001.1262.

Loconte, H., and Estes, J. R. (1989). Phylogenetic Systematics of Berberidaceae and Ranunculales (Magnoliidae). *Syst. Bot.* 14, 565–579. doi:10.2307/2419001.

Loehrer, M., Langenbach, C., Goellner, K., Conrath, U., and Schaffrath, U. (2008). Characterization of nonhost resistance of Arabidopsis to the Asian soybean rust. *Mol. Plant-Microbe Interact. MPMI* 21, 1421–1430. doi:10.1094/MPMI-21-11-1421.

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. O., and Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* 33, 6494–6506. doi:10.1093/nar/gki937.

Loughman, R., Jayasena, K., and Majewski, J. (2005). Yield loss and fungicide control of stem rust of wheat. *Aust. J. Agric. Res.* 56, 91–96. Available at: https://doi.org/10.1071/AR04126.

Lowe, T. M., and Chan, P. P. (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44, W54-57. doi:10.1093/nar/gkw413.

Luig, N. H. (1978). Mutation studies in *Puccinia graminis tritici*, Proceedings of the 5th International Wheat Genetics Symposium. *New Delhi Indian Soc. Genet. Plant Breed.* 1, 533–539.

Luig, N. H. (1985). "Epidemiology in Australia and New Zealand," in *Diseases, Distribution, Epidemiology, and Control* (Elsevier), 301–328.

Luig, N., and Watson, I. (1972). The Role of Wild and Cultivated Grasses in the Hybridization of Formae Speciales of *Puccinia graminis*. *Aust. J. Biol. Sci.* 25, 335. doi:10.1071/BI9720335.

Mago, R., Zhang, P., Vautrin, S., Šimková, H., Bansal, U., Luo, M.-C., et al. (2015). The wheat *Sr50* gene reveals rich diversity at a cereal disease resistance locus. *Nat. Plants* 1, 15186. doi:10.1038/nplants.2015.186.

Mao, G., Meng, X., Liu, Y., Zheng, Z., Chen, Z., and Zhang, S. (2011). Phosphorylation of a WRKY Transcription Factor by Two Pathogen-Responsive MAPKs Drives Phytoalexin Biosynthesis in *Arabidopsis*. *Plant Cell* 23, 1639–1653. doi:10.1105/tpc.111.084996.

Margulis, L., and Chapman, M. J. (2009). "Chapter Four - KINGDOM FUNGI," in *Kingdoms and Domains (Fourth Edition)*, eds. L. Margulis and M. J. Chapman (London: Academic Press), 379–409. doi:10.1016/B978-0-12-373621-5.00004-0.

Mellersh, D. G., and Heath, M. C. (2003). An Investigation into the Involvement of Defense Signaling Pathways in Components of the Nonhost Resistance of *Arabidopsis thaliana* to Rust Fungi Also Reveals a Model System for Studying Rust Fungal Compatibility. *Mol. Plant. Microbe Interact.* 16, 398–404. doi:10.1094/MPMI.2003.16.5.398.

Mettin, D., Bluthner, W. D., and Schlegel, G. (1973). Additional evidence on spontaneous 1B/1R wheat-rye substitutions and translocations. in *Proceedings of the fourth international wheat genetics symposium. Alien genetic material.* (University of Missouri.), 179–184.

Meyer, M., Burgin, L., Hort, M. C., Hodson, D. P., and Gilligan, C. A. (2017a). Large-Scale Atmospheric Dispersal Simulations Identify Likely Airborne Incursion Routes of Wheat Stem Rust Into Ethiopia. *Phytopathology* 107, 1175–1186. doi:10.1094/PHYTO-01-17-0035-FI.

Meyer, M., Cox, J. A., Hitchings, M. D. T., Burgin, L., Hort, M. C., Hodson, D. P., et al. (2017b). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nat. Plants* 3, 780–786. doi:10.1038/s41477-017-0017-5.

Miller, M. E., Zhang, Y., Omidvar, V., Sperschneider, J., Schwessinger, B., Raley, C., et al. (2018). De novo assembly and phasing of dikaryotic genomes from two isolates of *Puccinia coronata* f. sp. *avenae*, the causal agent of oat crown rust. *mBio* 9, e01650-17.

Milne, R. J., Dibley, K. E., and Lagudah, E. S. (2017). Yeast as a Heterologous System to Functionally Characterize a Multiple Rust Resistance Gene that Encodes a Hexose Transporter. *Methods Mol. Biol. Clifton NJ* 1659, 265–274. doi:mn.

Milne, R. J., Dibley, K. E., Schnippenkoetter, W., Mascher, M., Lui, A. C. W., Wang, L., et al. (2019). The Wheat *Lr67* Gene from the Sugar Transport Protein 13 Family Confers Multipathogen Resistance in Barley. *Plant Physiol.* 179, 1285–1297. doi:10.1104/pp.18.00945.

Milus, E. A., Kristensen, K., and Hovmøller, M. S. (2009). Evidence for Increased Aggressiveness in a Recent Widespread Strain of *Puccinia striiformis* f. sp. *tritici* Causing Stripe Rust of Wheat. *Phytopathology* 99, 89–94. doi:10.1094/PHYTO-99-1-0089.

Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Curr. Opin. Plant Biol.* 15, 349–357. doi:10.1016/j.pbi.2012.05.006.

Moore, J. W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., et al. (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 47, 1494–1498. doi:10.1038/ng.3439.

Mysore, K. S., and Ryu, C.-M. (2004). Nonhost resistance: how much do we know? *Trends Plant Sci.* 9, 97–104. doi:10.1016/j.tplants.2003.12.005.

Niks, R. E., Parlevliet, J. E., Lindhout, P., and Bai, Y. (2011). *Breeding crops with resistance to diseases and pests.* Wageningen Academic Publishers.

Nørholm, M. H. H., Nour-Eldin, H. H., Brodersen, P., Mundy, J., and Halkier, B. A. (2006). Expression of the *Arabidopsis* high-affinity hexose transporter STP13 correlates with programmed cell death. *FEBS Lett.* 580, 2381–2387. doi:10.1016/j.febslet.2006.03.064.

Oliver, R. P. (2014). A reassessment of the risk of rust fungi developing resistance to fungicides. *Pest Manag. Sci.* 70, 1641–1645. doi:10.1002/ps.3767.

Olivera, P. D., Newcomb, M., Flath, K., Sommerfeldt-Impe, N., Szabo, L. J., Carter, M., et al. (2017). Characterization of Puccinia graminis f. sp. tritici isolates derived from an unusual wheat stem rust outbreak in Germany in 2013. *Plant Pathol.* 66, 1258–1266.

Olivera, P., Newcomb, M., Szabo, L. J., Rouse, M., Johnson, J., Gale, S., et al. (2015). Phenotypic and Genotypic Characterization of Race TKTTF of *Puccinia graminis* f. sp. *tritici* that Caused a Wheat Stem Rust Epidemic in Southern Ethiopia in 2013–14. *Phytopathology* 105, 917–928. doi:10.1094/PHYTO-11-14-0302-FI.

Omidvar, V., Dugyala, S., Li, F., Rottschaefer, S. M., Miller, M. E., Ayliffe, M., et al. (2018). Detection of Race-Specific Resistance Against *Puccinia coronata* f. sp.

avenae in Brachypodium Species. *Phytopathology* 108, 1443–1454. doi:10.1094/PHYTO-03-18-0084-R.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290.

Pardey, P. G., Beddow, J. M., Kriticos, D. J., Hurley, T. M., Park, R. F., Duveiller, E., et al. (2013). Right-sizing stem-rust research. *Science* 340, 147–148. Available at: http://science.sciencemag.org/content/340/6129/147.short [Accessed October 23, 2016].

Park, R. F. (2007). Stem rust of wheat in Australia. *Aust. J. Agric. Res.* 58, 558–566. doi:10.1071/AR07117.

Park, R. F., Burdon, J. J., and Jahoor, A. (1999). Evidence for somatic hybridization in nature in *Puccinia recondita* f. sp. *tritici*, the leaf rust pathogen of wheat. *Mycol. Res.* 103, 715–723. doi:10.1017/S0953756298007631.

Park, R. F., and Wellings, C. R. (2012). Somatic Hybridization in the Uredinales. *Annu. Rev. Phytopathol.* 50, 219–239. doi:10.1146/annurev-phyto-072910-095405.

Patpour, M., Hovmøller, M. S., Justesen, A. F., Newcomb, M., Olivera, P., Jin, Y., et al. (2016). Emergence of Virulence to *SrTmp* in the Ug99 Race Group of Wheat Stem Rust, *Puccinia graminis* f. sp. *tritici*, in Africa. *Plant Dis.* 100, 522–522. doi:10.1094/PDIS-06-15-0668-PDN.

Periyannan, S., Milne, R. J., Figueroa, M., Lagudah, E. S., and Dodds, P. N. (2017). An overview of genetic rust resistance: From broad to specific mechanisms. *PLOS Pathog.* 13, e1006380. doi:10.1371/journal.ppat.1006380.

Periyannan, S., Moore, J., Ayliffe, M., Bansal, U., Wang, X., Huang, L., et al. (2013). The Gene *Sr33*, an Ortholog of Barley Mla Genes, Encodes Resistance to Wheat Stem Rust Race Ug99. *Science* 341, 786–788. doi:10.1126/science.1239028.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290.

Pretorius, Z. A., Pakendorf, K. W., Marais, G. F., Prins, R., and Komen, J. S. (2007). Challenges for sustainable cereal rust control in South Africa. *Aust. J. Agric. Res.* 58, 593–601.

Pretorius, Z. A., Singh, R. P., Wagoire, W. W., and Payne, T. S. (2000). Detection of Virulence to Wheat Stem Rust Resistance Gene *Sr31* in *Puccinia graminis* f. sp. *tritici* in Uganda. *Plant Dis.* 84, 203–203. doi:10.1094/PDIS.2000.84.2.203B.

Pretorius, Z. A., Szabo, L. J., Boshoff, W. H. P., Herselman, L., and Visser, B. (2012). First Report of a New TTKSF Race of Wheat Stem Rust (*Puccinia graminis* f. sp. *tritici*) in South Africa and Zimbabwe. *Plant Dis.* 96, 590–590. doi:10.1094/PDIS-12-11-1027-PDN.

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi:10.1093/bioinformatics/btq033.

Ravensdale, M., Nemri, A., Thrall, P. H., Ellis, J. G., and Dodds, P. N. (2011). Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. *Mol. Plant Pathol.* 12, 93–102. doi:10.1111/j.1364-3703.2010.00657.x.

Rawlings, N. D. (2016). Peptidase specificity from the substrate cleavage collection in the MEROPS database and a tool to measure cleavage site conservation. *Biochimie* 122, 5–30.

Rentsch, D., Laloi, M., Rouhara, I., Schmelzer, E., Delrot, S., and Frommer, W. B. (1995). *NTR1* encodes a high affinity oligopeptide transporter in *Arabidopsis*. *FEBS Lett.* 370, 264–268. doi:10.1016/0014-5793(95)00853-2.

Roelfs, A., and Bushnell, W. (1984). Race specificity and methods of study. *Cereal Rusts* 1, 131–164.

Roelfs, A. P. (1978). *Estimated losses caused by rust in small grain cereals in the United States, 1918-76*. Dept. of Agriculture, Agricultural Research Service.

Roelfs, A. P. (1982). Effects of Barberry eradication. *Plant Dis.* 66, 177.

Roelfs, A. P. (1992). *Rust diseases of wheat: concepts and methods of disease management*. CIMMYT.

Roelfs, A. P., and Martens, J. W. (1988). An international system of nomenclature for *Puccinia graminis* f. sp. *tritici*. *Phytopathology* 78, 526–533.

Rosso, M. G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol. Biol.* 53, 247–259.

Rushton, P. J., Torres, J. T., Parniske, M., Wernert, P., Hahlbrock, K., and Somssich, I. E. (1996). Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J.* 15, 5690–5700. doi:10.1002/j.1460-2075.1996.tb00953.x.

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behav. Ecol.* 17, 688–690.

Saari, E. E., and Prescott, J. M. (1985). “9 - World Distribution in Relation to Economic Losses,” in *Diseases, Distribution, Epidemiology, and Control*, eds. A. P. Roelfs and W. R. Bushnell (Academic Press), 259–298. doi:10.1016/B978-0-12-148402-6.50017-1.

Saintenac, C., Zhang, W., Salcedo, A., Rouse, M. N., Trick, H. N., Akhunov, E., et al. (2013). Identification of Wheat Gene *Sr35* That Confers Resistance to Ug99 Stem Rust Race Group. *Science* 341, 783–786. doi:10.1126/science.1239022.

Salcedo, A., Rutter, W., Wang, S., Akhunova, A., Bolus, S., Chao, S., et al. (2017). Variation in the *AvrSr35* gene determines *Sr35* resistance against wheat stem rust race Ug99. *Science* 358, 1604–1606. doi:10.1126/science.aao7294.

Savile, D. B. O. (1984). Taxonomy of the cereal rust fungi. *Cereal Rusts Orig. Specif. Struct. Physiol.* 1, 79–112.

Schliep, K. P. (2010). phangorn: phylogenetic analysis in R. *Bioinformatics* 27, 592–593.

Schulze-Lefert, P., and Panstruga, R. (2011). A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends Plant Sci.* 16, 117–125. doi:10.1016/j.tplants.2011.01.001.

Schwessinger, B., and Rathjen, J. P. (2017). “Extraction of high molecular weight DNA from fungal rust spores for long read sequencing,” in *Wheat Rust Diseases* (Springer), 49–57.

Schwessinger, B., Sperschneider, J., Cuddy, W. S., Garnica, D. P., Miller, M. E., Taylor, J. M., et al. (2018). A near-complete haplotype-phased genome of the dikaryotic wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici* reveals high interhaplotype diversity. *mBio* 9, e02275-17.

Schwessinger, B., and Zipfel, C. (2008). News from the frontline: recent insights into PAMP-triggered immunity in plants. *Curr. Opin. Plant Biol.* 11, 389–395. doi:10.1016/j.pbi.2008.06.001.

Senthil-Kumar, M., and Mysore, K. S. (2013). Nonhost Resistance Against Bacterial Pathogens: Retrospectives and Prospects. *Annu. Rev. Phytopathol.* 51, 407–427. doi:10.1146/annurev-phyto-082712-102319.

Shafiei, R., Hang, C., Kang, J.-G., and Loake, G. J. (2007). Identification of loci controlling non-host disease resistance in Arabidopsis against the leaf rust pathogen *Puccinia triticina*. *Mol. Plant Pathol.* 8, 773–784. doi:10.1111/j.1364-3703.2007.00431.x.

Sharma, S., and Prasada, R. (1969). Production of new races of *Puccinia graminis* var *tritici*. *Aust. J. Agric. Res.* 20, 981–985. Available at: https://doi.org/10.1071/AR9690981.

Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. (1988). Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Res.* 16, 7583–7600. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC338428/ [Accessed August 26, 2019].

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.

Singh, R. (1992). Association Between Gene *Lr34* for Leaf Rust Resistance and Leaf Tip Necrosis in Wheat. *Crop Sci.* 32, 874–878. doi:10.2135/cropsci1992.0011183X003200040008x.

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., et al. (2011). The Emergence of Ug99 Races of the Stem Rust Fungus is a Threat to World Wheat Production. *Annu. Rev. Phytopathol.* 49, 465–481. doi:10.1146/annurev-phyto-072910-095423.

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Njau, P., Wanyera, R., et al. (2008). "Will Stem Rust Destroy the World's Wheat Crop?," in *Advances in Agronomy* (Elsevier), 271–309. doi:10.1016/S0065-2113(08)00205-8.

Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., et al. (2015). Emergence and Spread of New Races of Wheat Stem Rust Fungus: Continued Threat to Food Security and Prospects of Genetic Control. *Phytopathology* 105, 872–884. doi:10.1094/PHYTO-01-15-0030-FI.

Singh, R. P., Singh, P. K., Rutkoski, J., Hodson, D. P., He, X. Y., Jorgensen, L. N., et al. (2016). Disease Impact on Wheat Yield Potential and Prospects of Genetic Control. *Annu. Rev. Phytopathol. Vol 54* 54, 303–322. doi:10.1146/annurev-phyto-080615-095835.

Slater, G. S. C., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6, 31. doi:10.1186/1471-2105-6-31.

Slewinski, T. L. (2011). Diverse Functional Roles of Monosaccharide Transporters and their Homologs in Vascular Plants: A Physiological Perspective. *Mol. Plant* 4, 641–662. doi:10.1093/mp/ssr051.

Solanki, S., Ameen, G., Borowicz, P., and Brueggeman, R. S. (2019). Shedding Light on Penetration of Cereal Host Stomata by Wheat Stem Rust Using Improved Methodology. *Sci. Rep.* 9. doi:10.1038/s41598-019-44280-6.

Somssich, I. E. (2007). "Networks of Transcriptional Regulation Underlying Plant Defense Responses Toward Phytopathogens," in *Annual Plant Reviews Volume 29:*

Regulation of Transcription in Plants (John Wiley & Sons, Ltd), 266–284. doi:10.1002/9780470988886.ch11.

Soreng, R. J., Peterson, P. M., Romaschenko, K., Davidse, G., Zuloaga, F. O., Judziewicz, E. J., et al. (2015). A worldwide phylogenetic classification of the Poaceae (Gramineae). *J. Syst. Evol.* 53, 117–137. doi:10.1111/jse.12150.

Spatafora, J. W., Aime, M. C., Grigoriev, I. V., Martin, F., Stajich, J. E., and Blackwell, M. (2017). The fungal tree of life: from molecular systematics to genome-scale phylogenies. *Microbiol. Spectr.* 5, 1–32.

Stajich, J. E., Berbee, M. L., Blackwell, M., Hibbett, D. S., James, T. Y., Spatafora, J. W., et al. (2009). Primer -- The Fungi. *Curr. Biol. CB* 19, R840–R845. doi:10.1016/j.cub.2009.07.004.

Stakman, E. C. (1947). Plant diseases are shifty enemies. *Sci. Prog. (New Haven)* 5, 235–279.

Stakman, E. C., and Harrar, J. G. (1958). Principles of plant pathology. *Soil Sci.* 85, 174.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Stanke, M., and Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 33, W465-467. doi:10.1093/nar/gki458.

Staples, R. C. (2001). Nutrients for a rust fungus: the role of haustoria. *Trends Plant Sci.* 6, 496–498. doi:10.1016/S1360-1385(01)02126-4.

Stein, M., Dittgen, J., Sánchez-Rodríguez, C., Hou, B.-H., Molina, A., Schulze-Lefert, P., et al. (2006). *Arabidopsis* PEN3/PDR8, an ATP Binding Cassette Transporter, Contributes to Nonhost Resistance to Inappropriate Pathogens That Enter by Direct Penetration. *Plant Cell* 18, 731–746. doi:10.1105/tpc.105.038372.

Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., Yu, G., et al. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34, 652–655. doi:10.1038/nbt.3543.

Stukenbrock, E. H. (2016). The role of hybridization in the evolution and emergence of new fungal plant pathogens. *Phytopathology* 106, 104–112.

Sucher, J., Boni, R., Yang, P., Rogowsky, P., Büchner, H., Kastner, C., et al. (2017). The durable wheat disease resistance gene *Lr34* confers common rust and northern corn leaf blight resistance in maize. *Plant Biotechnol. J.* 15, 489–496. doi:10.1111/pbi.12647.

Terefe, T., Pretorius, Z. A., Visser, B., and Boshoff, W. H. P. (2018). First Report of *Puccinia graminis* f. sp. *tritici* Race PTKSK, a Variant of Wheat Stem Rust Race Ug99 in South Africa. *Plant Dis.* doi:10.1094/PDIS-11-18-1911-PDN.

Testa, A. C., Hane, J. K., Ellwood, S. R., and Oliver, R. P. (2015). CodingQuarry: highly accurate hidden Markov model gene prediction in fungal genomes using RNA-seq transcripts. *BMC Genomics* 16, 170. doi:10.1186/s12864-015-1344-4.

Toruño, T. Y., Stergiopoulos, I., and Coaker, G. (2016). Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. *Annu. Rev. Phytopathol.* 54, 419–441. doi:10.1146/annurev-phyto-080615-100204.

Tripathi, P., Rabara, R. C., Langum, T. J., Boken, A. K., Rushton, D. L., Boomsma, D. D., et al. (2012). The WRKY transcription factor family in *Brachypodium distachyon*. *BMC Genomics* 13, 270. doi:10.1186/1471-2164-13-270.

Tzfira, T., Li, J., Lacroix, B., and Citovsky, V. (2004). Agrobacterium T-DNA integration: molecules and models. *Trends Genet.* 20, 375–383. doi:10.1016/j.tig.2004.06.004.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40, e115. doi:10.1093/nar/gks596.

Upadhyaya, N. M., Garnica, D. P., Karaoglu, H., Sperschneider, J., Nemri, A., Xu, B., et al. (2015). Comparative genomics of Australian isolates of the wheat stem rust pathogen *Puccinia graminis* f. sp. *tritici* reveals extensive polymorphism in candidate effector genes. *Front. Plant Sci.* 5, 759. doi:10.3389/fpls.2014.00759.

van Schie, C. C. N., and Takken, F. L. W. (2014). Susceptibility Genes 101: How to Be a Good Host. *Annu. Rev. Phytopathol.* 52, 551–581. doi:10.1146/annurev-phyto-102313-045854.

Vanegas, C. D. G., Garvin, D. F., and Kolmer, J. A. (2008). Genetics of stem rust resistance in the spring wheat cultivar Thatcher and the enhancement of stem rust resistance by *Lr34*. *Euphytica* 159, 391–401. doi:10.1007/s10681-007-9541-0.

Viljanen, S., and Cromey, M. (2002). Pathways of entry and spread of rust pathogens: implications for New Zealand's biosecurity. N Z Plant Prot. *N. Z. Plant Prot.* 55.

Visser, B., Herselman, L., Park, R. F., Karaoglu, H., Bender, C. M., and Pretorius, Z. A. (2011). Characterization of two new *Puccinia graminis* f. sp. *tritici* races within the Ug99 lineage in South Africa. *Euphytica* 179, E119-127.

Visser, B., Meyer, M., Park, R. F., Gilligan, C. A., Burgin, L. E., Hort, M. C., et al. (2019). Microsatellite Analysis and Urediniospore Dispersal Simulations Support the

Movement of *Puccinia graminis* f. sp. *tritici* from Southern Africa to Australia. *Phytopathology* 109, 133–144. doi:10.1094/PHYTO-04-18-0110-R.

Voegele, R. T., and Mendgen, K. W. (2011). Nutrient uptake in rust fungi: how sweet is parasitic life? *Euphytica* 179, 41–55. doi:10.1007/s10681-011-0358-5.

Vogel, J., and Hill, T. (2008). High-efficiency Agrobacterium-mediated transformation of Brachypodium distachyon inbred line Bd21-3. *Plant Cell Rep.* 27, 471–478. doi:10.1007/s00299-007-0472-y.

Wang, X., and McCallum, B. (2009). Fusion Body Formation, Germ Tube Anastomosis, and Nuclear Migration During the Germination of Urediniospores of the Wheat Leaf Rust Fungus, *Puccinia triticina*. *Phytopathology* 99, 1355–1364. doi:10.1094/PHYTO-99-12-1355.

Wanyera, R., Kinyua, M. G., Jin, Y., and Singh, R. P. (2006). The Spread of Stem Rust Caused by *Puccinia graminis* f. sp. *tritici* , with Virulence on *Sr31* in Wheat in Eastern Africa. *Plant Dis.* 90, 113–113. doi:10.1094/PD-90-0113A.

Wanyera, R., Macharia, J. K., Kilonzo, S. M., and Kamundia, J. W. (2009). Foliar fungicides to control wheat stem rust, race TTKS (Ug99), in Kenya. *Plant Dis.* 93, 929–932.

Watson, I. (1957). Further studies on the production of new races from mixtures of races of *Puccinia graminis* var. *tritici* on wheat seedlings. *Phytopathology* 47, 510–512.

Watson, I. A., and Luig, N. H. (1958). Somatic hybridization in *Puccinia graminis* var *tritici*. *Proc. Linn. Soc. New South Wales* 83, 190–195.

Watson, I. A., and Sousa, C. N. A. de (1983). Long distance transport of spores of *Puccinia graminis tritici* in the southern hemisphere. *Proc. Linn. Soc. New South Wales* 106, 311–321.

Wellings, C. (2012a). Stripe rust resistance in triticale in Australia: pathogen change exposes genetic vulnerability. in *Proceedings of the 13th International Cereal Rusts and Powdery Mildews Conference* (Beijing, China: China Agricultural Sciences and Technology Press), 38–39.

Wellings, C. R. (2012b). Resistance to stripe rust in wheat: pathogen biology driving resistance breeding. *Dis. Resist. Wheat Sharma Ed*, 63–83.

Wellings, C.R., C. R., McIntosh, R. A., and Walker, J. (1987). *Puccinia striiformis* f. sp. *tritici* in Eastern Australia possible means of entry and implications for plant quarantine. *Plant Pathol.* 36, 239–241. doi:10.1111/j.1365-3059.1987.tb02230.x.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wieczorke, R., Krampe, S., Weierstall, T., Freidel, K., Hollenberg, C. P., and Boles, E. (1999). Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett.* 464, 123–128. Available at: http://onlinelibrary.wiley.com/doi/10.1016/S0014-5793(99)01698-1/full [Accessed October 23, 2016].

Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., et al. (2015). HiCUP: pipeline for mapping and processing Hi-C data. *F1000Research* 4.

Wulff, B. B., Horvath, D. M., and Ward, E. R. (2011). Improving immunity in crops: new tactics in an old game. *Curr. Opin. Plant Biol.* 14, 468–476. doi:10.1016/j.pbi.2011.04.002.

Yamada, K., Saijo, Y., Nakagami, H., and Takano, Y. (2016). Regulation of sugar transporter activity for antibacterial defense in Arabidopsis. *Science* 354, 1427–1430. doi:10.1126/science.aah5692.

Yin, C., Chen, X., Wang, X., Han, Q., Kang, Z., and Hulbert, S. H. (2009). Generation and analysis of expression sequence tags from haustoria of the wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici*. *BMC Genomics* 10, 1–9. doi:10.1186/1471-2164-10-626.

Yin, T., Cook, D., and Lawrence, M. (2012a). ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol.* 13, R77. doi:10.1186/gb-2012-13-8-r77.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012b). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451.

Yirgou, D., and Caldwell, R. M. (1963). Stomatal Penetration of Wheat Seedlings by Stem and Leaf Rust: Effect of Light and Carbon Dioxide. *Science* 141, 272–273. doi:10.1126/science.141.3577.272.

Yu, Q., Tang, C., and Kuo, J. (2000). A critical review on methods to measure apoplastic pH in plants. *Plant Soil* 219, 29–40. doi:10.1023/A:1004724610550.

Yun, B.-W., Atkinson, H. A., Gaborit, C., Greenland, A., Read, N. D., Pallas, J. A., et al. (2003). Loss of actin cytoskeletal function and EDS1 activity, in combination, severely compromises non-host resistance in *Arabidopsis* against wheat powdery mildew. *Plant J. Cell Mol. Biol.* 34, 768–777. doi:10.1046/j.1365-313x.2003.01773.x.

Zambino, P. J., Kubelik, A. R., and Szabo, L. J. (2000). Gene action and linkage of avirulence genes to DNA markers in the rust fungus *Puccinia graminis*. *Phytopathology* 90, 819–826. Available at: http://apsjournals.apsnet.org/doi/abs/10.1094/PHYTO.2000.90.8.819 [Accessed December 11, 2014].

Zeller, F. J. (1973). 1B/1R wheat-rye chromosome substitutions and translocations. *Proc. Fourth Int. Wheat Genet. Symp. Alien Genet. Mater.*, 209–221. Available at: https://www.cabdirect.org/cabdirect/abstract/19751625251 [Accessed September 17, 2019].

Zellerhoff, N., Himmelbach, A., Dong, W., Bieri, S., Schaffrath, U., and Schweizer, P. (2010). Nonhost Resistance of Barley to Different Fungal Pathogens Is Associated with Largely Distinct, Quantitative Transcriptional Responses. *Plant Physiol.* 152, 2053–2066. doi:10.1104/pp.109.151829.

Zhang, H., Wang, C., Cheng, Y., Wang, X., Li, F., Han, Q., et al. (2011). Histological and molecular studies of the non-host interaction between wheat and *Uromyces fabae*. *Planta* 234, 979–991. doi:10.1007/s00425-011-1453-5.

Zhang, W., Chen, S., Abate, Z., Nirmala, J., Rouse, M. N., and Dubcovsky, J. (2017). Identification and characterization of *Sr13*, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. *Proc. Natl. Acad. Sci.* 114, E9483–E9492. doi:10.1073/pnas.1706277114.

Zhao, B., Lin, X., Poland, J., Trick, H., Leach, J., and Hulbert, S. (2005). A maize resistance gene functions against bacterial streak disease in rice. *Proc. Natl. Acad. Sci.* 102, 15383–15388. doi:10.1073/pnas.0503023102.

Zheng, Z., Qamar, S. A., Chen, Z., and Mengiste, T. (2006). Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J.* 48, 592–605. doi:10.1111/j.1365-313X.2006.02901.x.

Zipfel, C. (2008). Pattern-recognition receptors in plant innate immunity. *Curr. Opin. Immunol.* 20, 10–16. doi:10.1016/j.coi.2007.11.003.

Appendix

Appendix 1: Supplementary Materials of Chapter 2

Supplemental Table 2. Summary statistics for SMRT sequencing and raw read metrics.

	21-0	Ug99	
	RSII	RSII	Sequel
Number of SMRT cells	17	4	5
Total Bases (Gb)	17.4	2.48	19.66
Number of Reads	1,248,195	317,864	2,634,315
Mean Subread Length (bp)	10,239	7,790	7,591
N_{50} Subread Length (bp)	16,438	12,080	13,550

Supplementary Table 3.
Summary statistics of Illumina sequencing of Pgt isolates in the Ug99 lineage.

Isolate (Pathotype)	150 bp Paired-End Reads	Yield (Mbp)	Mean Quality Score
UVPgt55 (TTKSF)	40,048,658	12,094	31.53
UVPgt59 (TTKSP)	27,204,289	8,216	31.14
UVPgt60 (PTKST)	36,665,018	11,073	31.56
UVPgt61 (TTKSF)*	36,605,359	11,055	31.57
Ug99 (TTKSK)	35,674,381	10,773	31.48

* Virulent on resistance gene *Sr9h*

Supplementary Table 4.

Assembly metrics and quality analysis

Parameters	21-0	Ug99
No. of contigs	410	514
No. of contigs ≥ 50,000 bp	249	333
Total length (Mbp)	176.9	176.0
Total length ≥50000bp (Mbp)	171.8	170.4
Size of Largest contig (Mbp)	5.96	4.40
N_{50} (Mbp)	1.26	0.97
GC (%)	43.5	43.5
No. of contigs with telomeres	69	26
% of complete BUSCOs	95.8	95.6
% single-copy BUSCOs	8.3	8.9
% duplicated BUSCOs	87.5	86.7
% of fragmented BUSCOs	1.9	1.9
% of missing BUSCOs	2.3	2.5
No. of Bins	44	62
Total length in bins (Mbp)	169	165
No. contigs in bins	225	276

Supplementary Table 6. Intra and inter-isolate sequence comparison of the *AvrSr50* chromosome haplotypes in Ug99 and *Pgt*21-0.

	Sequence similarity							Structural variation (SV)		
									Total size of variants	
Isolate comparison[a]	**Bases aligned (%)**	**Average identity of alignment blocks (%)**	**Overall identity[b] (%)**	**Divergence of aligned blocks (%)**	**Total SNPs**	**SNPs/kbp**	**Indels**	**Number of variants > 50bp**	**Mbp**	**% of chromosome**
Ug99 A vs *Pgt*21-0 A	99.8	99.93	99.73	0.07	307	0.10	820	29	0.17	2.56
Ug99 A vs Ug99 C	70.82	95.12	67.36	4.88	52,839	25.28	34,563	167	1.3	22.03
21-0 A vs *Pgt*21-0 B	73.12	95.50	69.83	4.50	57,463	22.03	37,070	190	1	14.03
Ug99 C vs *Pgt*21-0 B	78.36	97.27	76.22	2.73	33,655	14.56	21,766	137	1.09	16.74
Ug99 A vs *Pgt*21-0 B	71.87	94.98	68.26	5.02	55,310	26.07	35,314	187	1.25	19.2
21-0 A vs Ug99 C	64.26	95.11	61.12	4.89	54,593	23.82	35,340	150	0.97	14.8

[a] First listed isolate served as reference and second listed isolate served as query for the analysis.

[b] Overall identity is average identity of alignment block multiplied by the proportion of bases aligned

Supplementary Table 9. Intra- and inter-isolate sequence comparison of entire haplotypes in Ug99 and *Pgt*21-0.

	Sequence similarity							Structural variation (SV)		
									Total size of variants	
Isolate comparison[a]	**Bases aligned (%)**	**Average identity of alignment blocks (%)**	**Overall identity**[b] **(%)**	**Divergence of aligned blocks (%)**	**Total SNPs**	**SNPs/Kbp**	**Indels**	**Number of variants >50bp**	**Mbp**	**% of genome**
21-0 A vs Ug99 A	99.64	99.92	99.56	0.08	9,275	0.10	24,835	491	0.82	0.46
Ug99 A vs Ug99 C	91.52	95.92	87.79	4.08	1,367,911	17.73	851,465	2,571	13.69	7.88
21-0 A vs *Pgt*21-0 B	91.38	97.60	87.55	2.40	1,418,591	17.71	877,814	2,696	15.01	8.56
21-0 B vs Ug99 C	93.44	95.82	91.20	4.18	876,653	11.13	572,042	1,910	11.50	6.69
Ug99 A vs *Pgt*21-0 B	91.52	95.90	87.69	4.10	1,414,244	17.63	877,352	2,648	14.69	8.29
21-0 A vs Ug99 C	91.54	95.81	87.79	4.19	1,371,178	17.77	851,247	2,585	13.88	8.08

[a] First listed isolate served as reference and second listed isolate served as query for the analysis.

[b] Overall identity is average identity of alignment block multiplied by the proportion of bases aligned.

Supplementary Table 11. Chromosomes sizes in *Pgt*21-0

Chromosome	size in A haplotype (bp)	size in B haplotype (bp)
1	6,156,315	6,527,486
2	6,062,178	6,110,382
3	6,034,412	4,933,094
4	5,966,401	6,360,166
5	5,557,100	7,276,977
6	5,553,668	5,248,565
7	5,183,406	5,503,882
8	5,112,795	2,821,965
9	4,787,417	5,140,183
10	4,647,647	4,889,217
11	4,639,132	4,947,173
12	3,976,497	3,939,087
13	3,569,361	3,304,927
14	3,567,101	3,561,970
15	3,495,074	3,444,174
16	3,430,011	5,891,779
17	3,317,526	2,935,361
18	2,873,293	3,063,918
total	83,929,334	85,900,306
	total size A and B	169,829,640

Supplementary Table 13. Summary of gene annotation

	21-0	Ug99
No. of genes including tRNAs	37,061	37,394
No. of protein coding genes	36,319	36,659
haplotype A	18,225 (17,786)*	18,593
haplotype B or C	17,919 (17,718)*	17,621
% of genome covered by genes	34.9	33.6
Mean gene length (bp)	1,667	1,586
No. of secreted protein genes	6,180	6,120
haplotype A	3,099 (3,071)*	3,212
haplotype B or C	3,063 (3,046)*	2,857

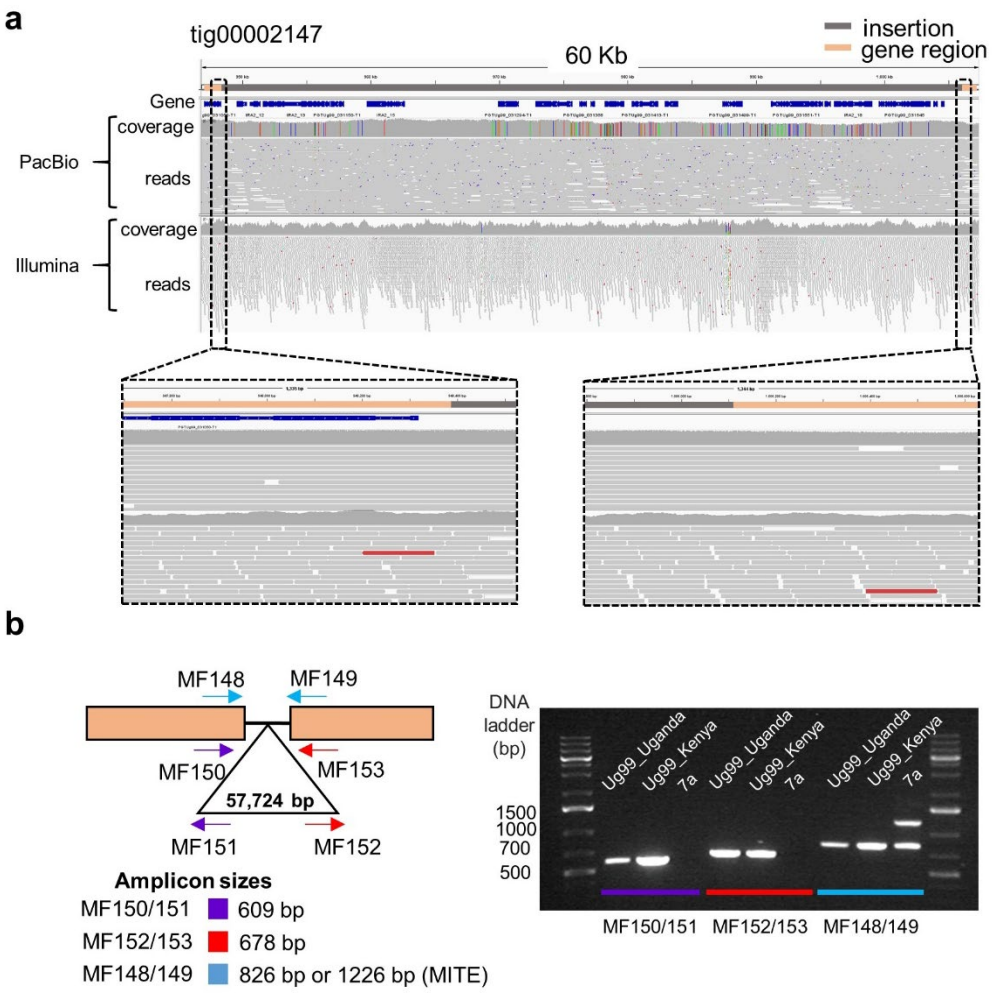
*No. of predicted genes in contigs assigned to chromosomes

Supplementary Table 14. Shared and unique gene content between *Pgt* haplotypes

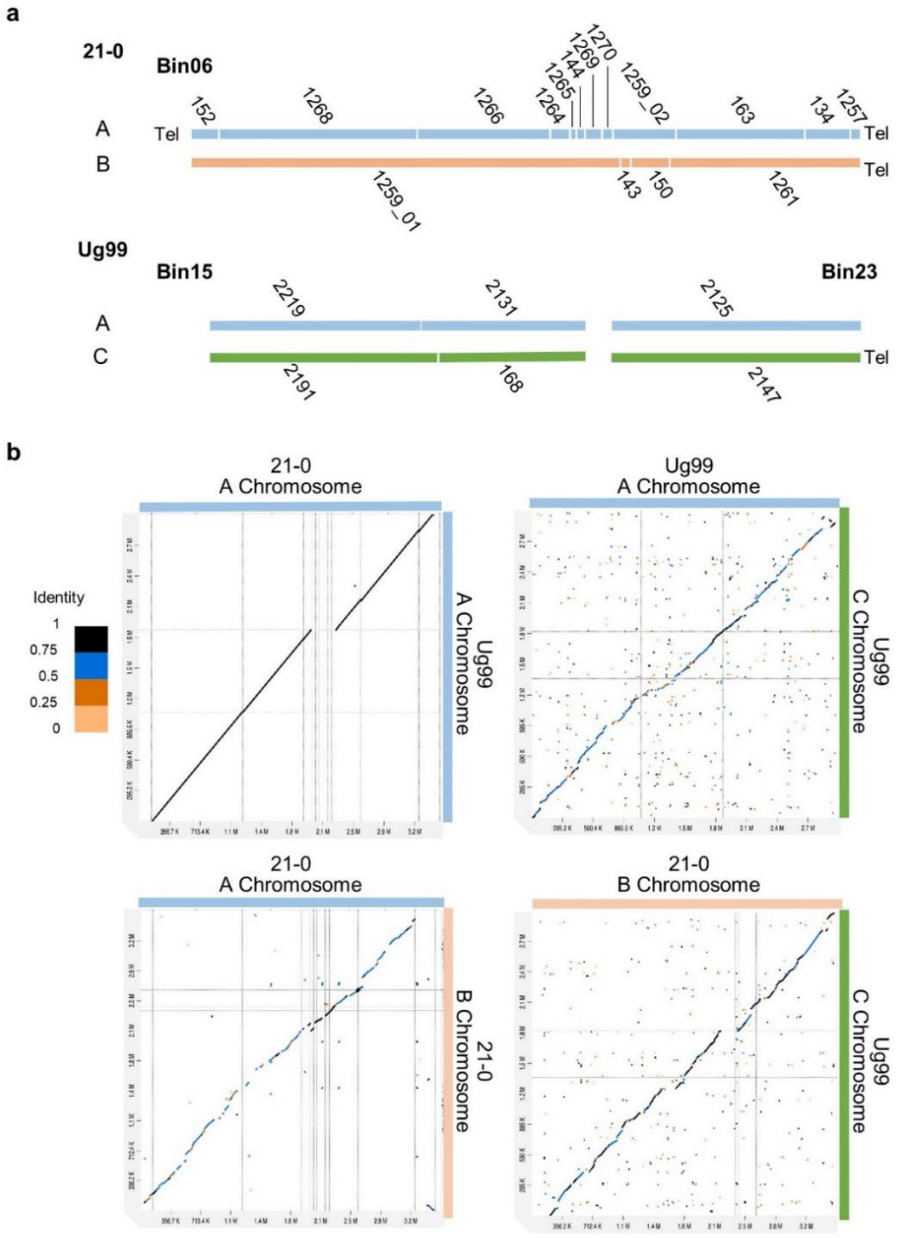
	Haplotypes compared			Haplotypes compared		
	21-0 A	21-0 B	Ug99 C	Ug99 A	21-0 B	Ug99 C
Unique genes	3,369 (18%)	2,668 (15%)	2,950 (17%)	3,529 (19%)	2,774 (15%)	2,950 (17%)
genes shared with one other haplotype	2,492 (14%)	3,165 (18%)	2,543 (14%)	2,976 (16%)	2,664 (15%)	2,827 (15%)
genes shared in three haplotypes	12,364 (68%)	12,086 (67%)	12,128 (69%)	12,088 (65%)	12,481 (70%)	12,082 (69%)
total genes	18,225	17,919	17,621	18,593	17,919	17,621

Supplementary Table 15. List of primer sequences to amplify flanking and internal regions of the 57 kbp-insert in *AvrSr35*.

Primer ID	Sequence 5'-3'	Amplicon size
MF148 (Forward)	TGCCAAAGTACAAATAGATGACCG	826 bp (with MITE sequence, 1,226 bp)
MF149 (Reverse)	AGATCTTTGAGGTGCTCCCC	
MF150 (Forward)	AGACAGTGTGAAATCAAGTACGT	609 bp
MF151 (Reverse)	CTCATGACAAGGGGCAGGG	
MF152 (Forward)	GCCCTTCAACATTCAGCCTC	678 bp
MF153 (Reverse)	GAGGTGCTCCCCAGGTATTA	

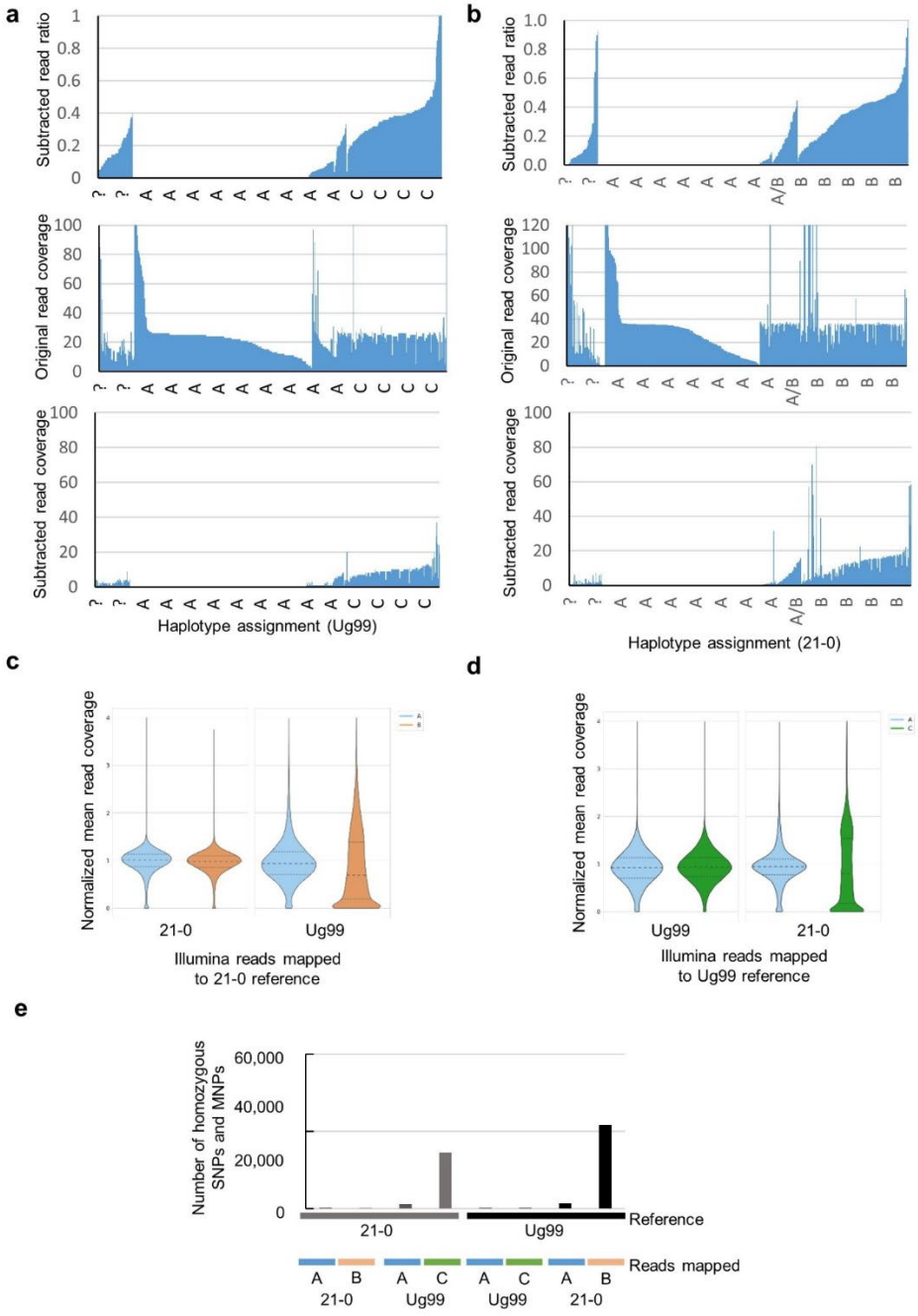


Supplemental Figure 1. Presence of a 57 kbp-insert in one allele of AvrSr35 in Ug99. **a,** Genome browser view of a 60 kbp genomic region in haplotype C of Ug99. The top bar shows the *AvrSr35* coding sequences (orange) flanking a 57 kbp-insert (grey). Annotated gene models (blue) are shown below. The following tracks show the read coverage graph and the alignments of Ug99 reads mapped to this region. Zoomed-in areas (boxed) show read mapping across the junction between the *AvrSr35* coding sequence and the 5' and 3' ends of the inserted sequence. **b,** Validation of 57 kbp-insert in *AvrSr35* of Ug99 isolates via PCR amplification. The positions of primers on the *AvrSr35* gene (orange boxes) and insertion (triangle) are shown along with the predicted amplicon sizes. PCR amplification products from the original Ug99 isolate (Ug99_Uganda), the Kenyan Ug99 isolate 04KEN156/04 (Ug99_Kenya) and the isolate CRL 75-36-700 (7a). Note that 7a is heterozygous for a wildtype allele of *AvrSr35* and a virulence allele containing a 400bp MITE insertion.



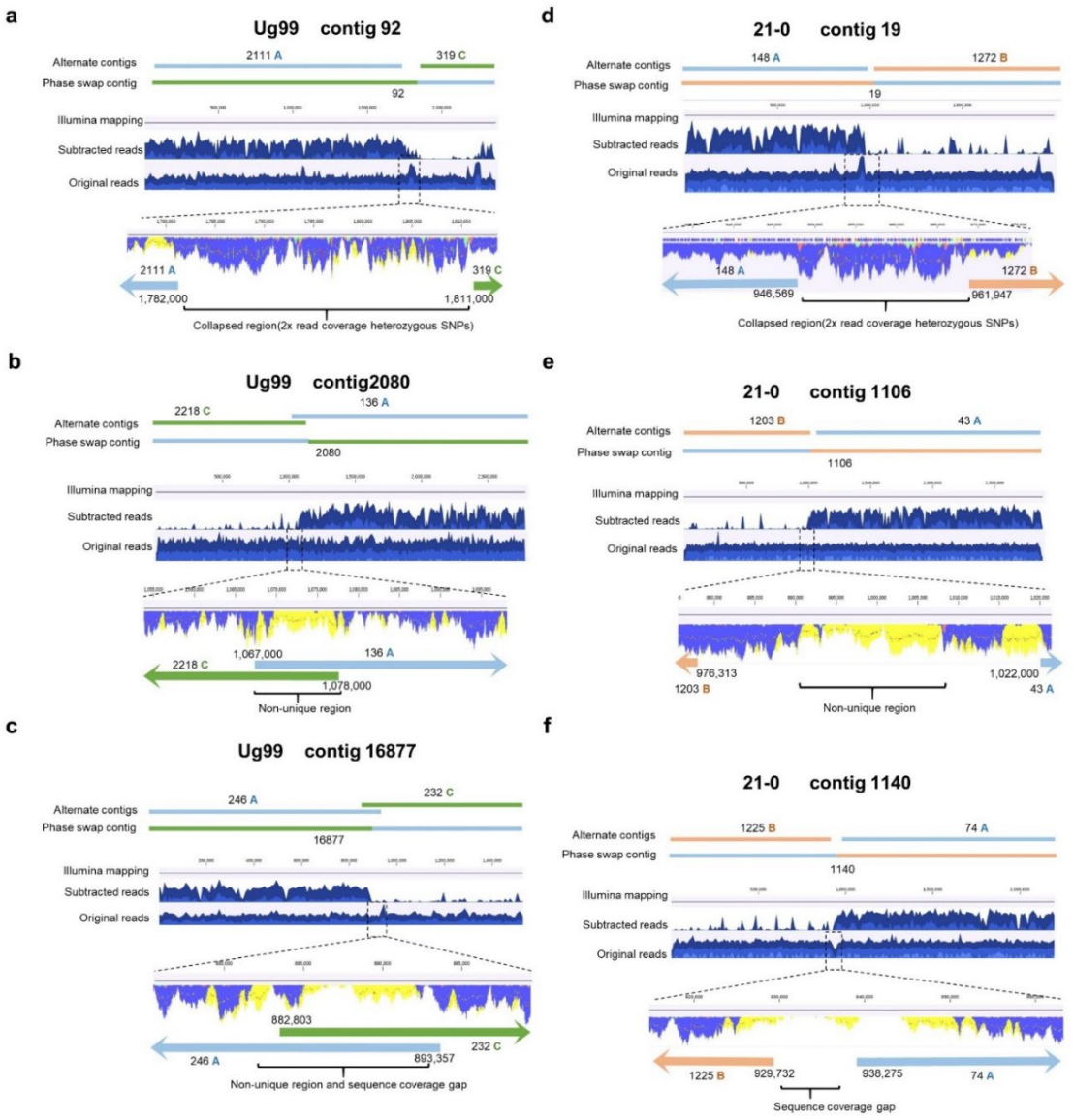
Supplementary Figure 2. One of the homologous chromosomes containing *AvrSr50* and *AvrSr35* loci is nearly identical in *Pgt*21-0 and Ug99.

a, Schematic representation of the alignment of contigs in *Pgt*21-0 and Ug99 derived from the homologous chromosomes. Contig IDs are indicated as numbers and presence of telomeres as "Tel". The *Pgt*21-0 contigs were assembled as Bin06 and contain telomeres at both ends indicating that a full chromosome was represented. The homologous contigs from Ug99 were present in two bins (Bin15 and Bin23). Contigs are coloured according to haplotype designation; A (light blue); B (orange); C (green). **b,** Dot plots of alignments between the homologous chromosomes of each haplotype, indicated by coloured bars at the top and right. X- and y-axes represent nucleotide positions. Colour key indicates sequence identity fraction for all dot plots (maximum identity score =1).



Supplementary Figure 3. Haplotype-specific read mapping and SNP calling validates the close identity of haplotype A in *Pgt* Ug99 and *Pgt*21-0.
a, Graphs of read depth ratio (top), and read depth of original (middle) and subtracted (bottom) Illumina reads (y-axis) from Ug99 mapped to Ug99 contigs listed by haplotype assignment A, C, A/C chimeras or unassigned (?) (x-axis). **b**, Graphs of read depth ratio (top), and read depth of original (middle) and subtracted (bottom) Illumina reads (y-axis) from *Pgt*21-0 mapped to *Pgt*21-0 contigs listed by haplotype assignment A, B, A/B

chimeras or unassigned (?) (x-axis). Data included in supplementary table 7. **c,** Violin plots for the distribution of read coverage for haplotype A (blue) and B (orange) after mapping Illumina reads from Ug99 or *Pgt*21-0 to the *Pgt*21-0 assembly. **d,** Violin plots for the distribution of read coverage for haplotype A (blue) and C (green) after mapping Illumina reads from Ug99 or *Pgt*21-0 to the Ug99 assembly. For **c** and **d** y-axis depicts genome coverage calculated in 1 kb sliding windows and normalized to the mean of coverage of each haplotype. Genome coverage shows a normal distribution for self-mapping. Read cross-mapping also shows a normal distribution for haplotype A of Ug99 and *Pgt*21-0 which indicates high sequence similarity. In contrast, a skewed distribution to low genome coverage occurs in the B and C haplotype comparison due to high sequence divergence. **e,** Numbers of homozygous SNPs and MNPs called for various extracted Illumina read sets mapped against the *Pgt*21-0 (grey) or Ug99 (black) reference genome assemblies. Illumina reads were first mapped at high stringency to the corresponding reference genome and then uniquely mapped reads from each haplotype were extracted and used for variant calling. The low number of SNPs detected in the inter isolate comparisons of haplotype A in contrast to the high number of SNPs identified in the B or C haplotype, supports the close identity of A haplotypes in both isolates.

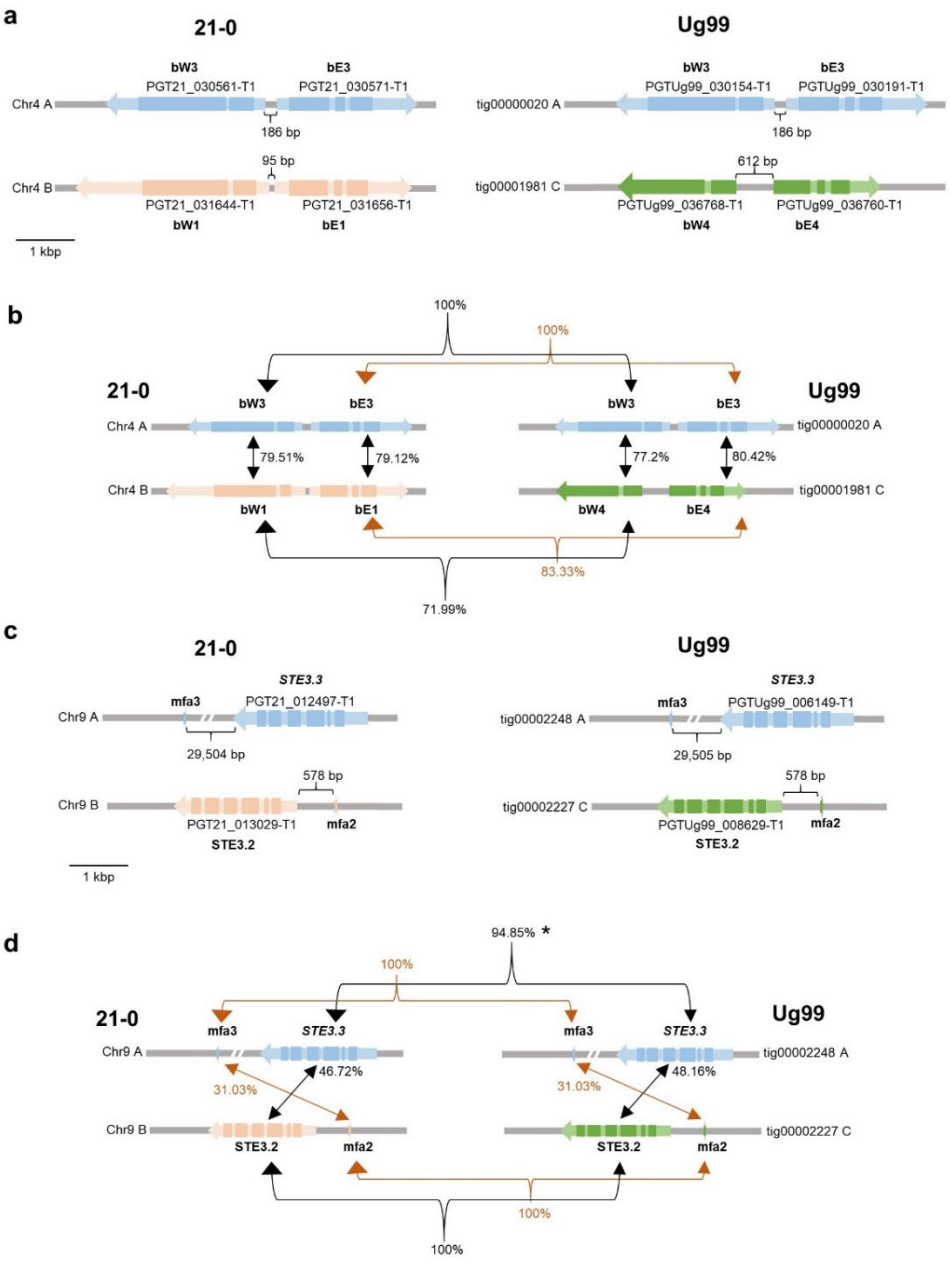


Supplementary Figure 4. Examples illustrating the detection and manual curation of phase swap contigs in the *Pgt*21-0 and Ug99 genome assemblies.

a to **f,** The top of each figure shows chimeric contigs and alternate contigs colour-coded according to haplotype assignment. The next two tracks show read coverage graphs of subtracted reads and original reads across the phase swap contigs visualized in CLC Genomics Workbench browser (see read subtraction procedure Fig. 3). Zoomed in regions (dotted boxes) show coverage graphs for the phase swap junction regions. Coloured bars indicate SNP frequencies in the underlying reads, and yellow shading indicates non-uniquely mapped reads. Coloured arrows at the bottom shows alignment positions of the alternate contigs to this region with the endpoint coordinates indicated. These examples illustrate scenarios indicative of assembly errors due to collapsed assembly regions showing double overage with heterozygous SNPs (**a, d**), non-uniquely mapped repeats (**b, e**) or coverage gaps after Illumina read mapping (**c, f**).

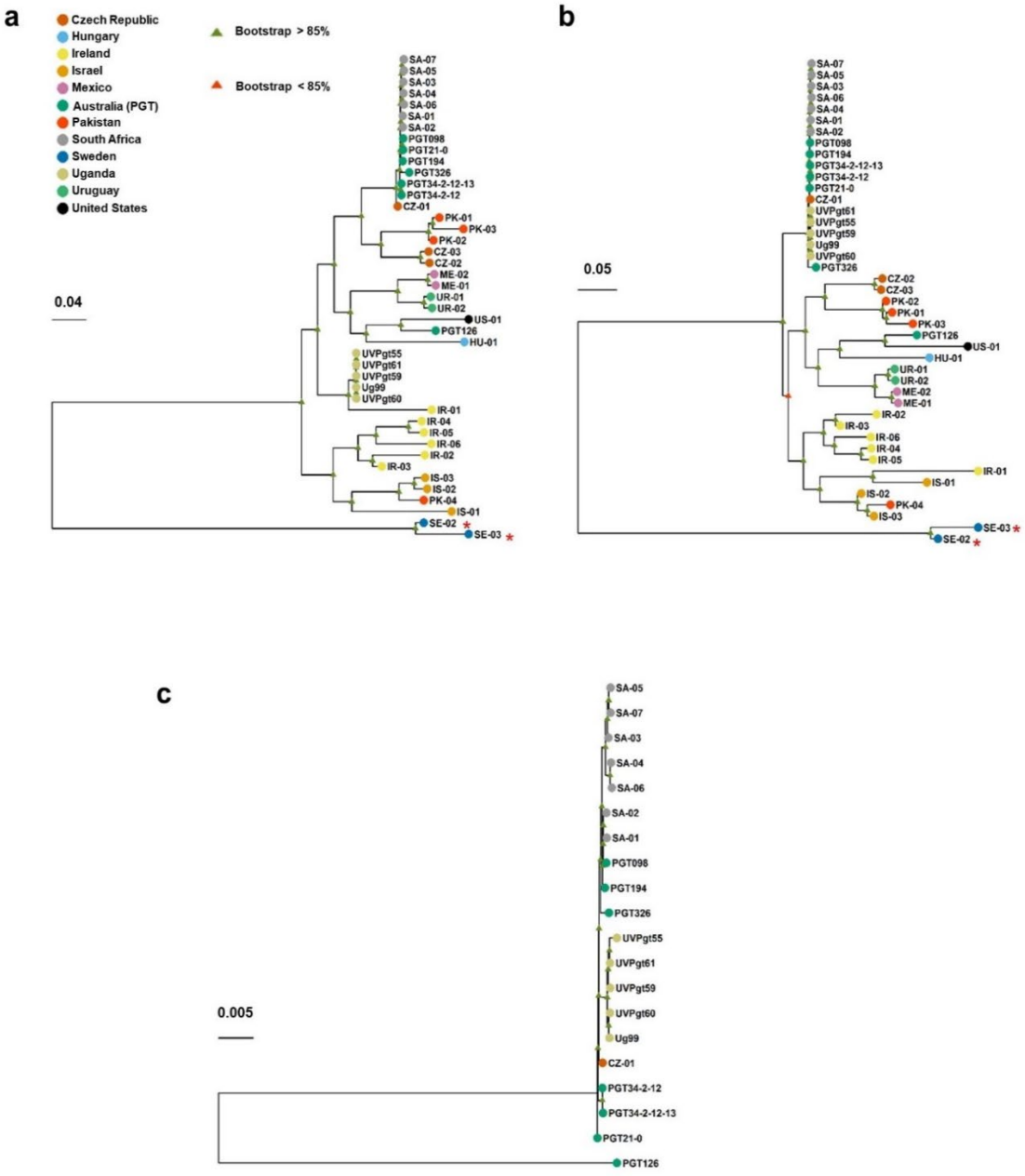
Supplementary Figure 5. Gene and repeat density plots for homologous chromosomes in haplotypes A and B of *Pgt*21-0.

Top two tracks show density of genes encoding non-secreted (black) or secreted proteins (red) along the chromosomes. Bottom graph shows density of repeat elements (blue). Positions of *bE/bW, STE3.2*, *STE 3.3*, *AvrSr50* and *AvrSr35* genes are indicated.



Supplementary Figure 6. Structure of mating type loci in *Pgt*21-0 and Ug99.
a, The gene transcripts and orientations of the *bE*/*bW* genes from the *b* mating-type locus (chromosome 4) are depicted by light coloured arrows and coding sequences by the darker boxes. Colour coding represents the three haplotypes (A=blue, B=orange, C=green). The distances between predicted gene models are shown. The *bW1*/*bE1* allele is identical to the *bE1*/*bW1* allele previously identified in a North American isolate 75-36-700-3. The *bE2/bW2* allele from 75-36-700-3 was not present in either isolate, which instead contained two additional novel alleles, *bE3*/*bW3* and *bE4*/*bW4*, indicating that this locus is multi-allelic in *Pgt*. **b,** Percentage amino acid identity between predicted proteins encoded by *bE* and *bW* alleles within and between *Pgt* isolates. **c,** Arrangement of the pheromone peptide

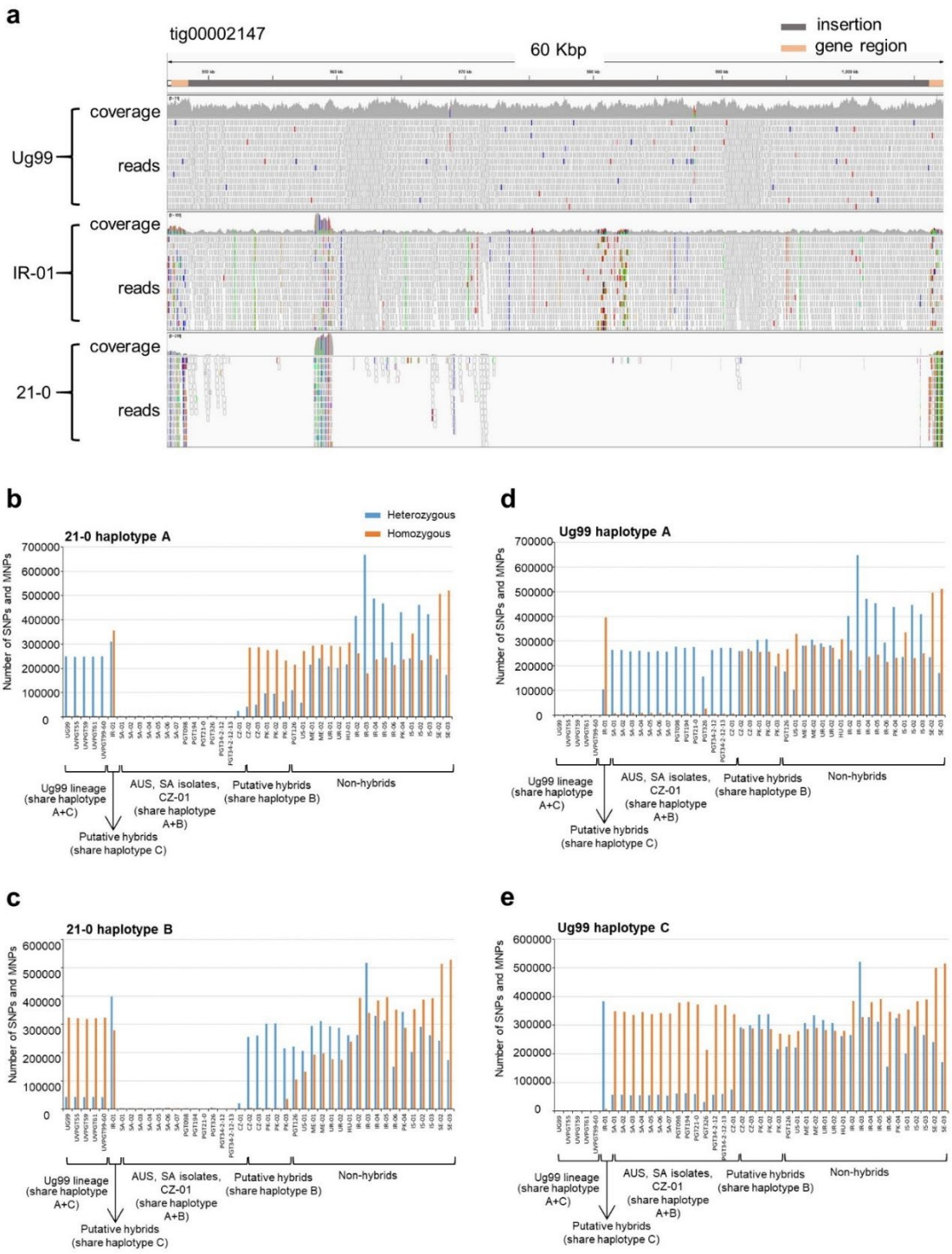
encoding genes (*mfa2* or *mfa3*) and pheromone mating factor receptors (*STE3.2* and *STE3.3*) at the predicted *a* mating type locus. The two alleles of the *a* locus on chromosome 9 contain either the *STE3.2* (B and C haplotypes) or *STE3.3* (A haplotype) genes from CRL 75-36-700-3 in both isolates, consistent with a binary recognition system. **d,** Percentage amino acid identity between pheromone peptide and receptor alleles within and between *Pgt* isolates. The *STE3.3* allele in Ug99 is identical to that in *Pgt*21-0 except for a 1 bp deletion causing a frameshift and replacement of the last 48 amino acids by an unrelated 24 amino acid sequence resulting in reduced amino acid identity (*).



Supplementary Figure 7. Phylogenetic analysis of *Pgt* isolates from diverse countries of origin.

a, Dendrogram inferred using biallelic SNPs detected against the complete diploid genome assembly of Ug99. **b,** Dendrogram inferred from SNPs detected in haplotype A of Ug99. **c,** Dendrogram inferred using SNPs in haplotype A of *Pgt*21-0 for the South African, Australian and Ug99 lineage isolates that share the A haploid genome, with *Pgt*126

included as an outgroup. Colour key in panel **b** indicates country of origin for all dendrograms. Scale bar indicates number of nucleotide substitutions per site. Red asterisks indicate *P. graminis* f. sp. *avenae* isolates.



Supplementary Figure 8. Putative *Pgt* hybrids that share the B or C haplotypes of *Pgt*21-0 and Ug99, respectively.

a, Genome browser view in IGV of a 60 kbp genomic region in haplotype C of Ug99. The top bar shows the *AvrSr35* coding sequences (orange) flanking a 57 kbp-insert (grey).

Following tracks illustrate coverage and Illumina read alignments of Ug99, IR-01, and *Pgt*21-0. In contrast to *Pgt*21-0, the genome of IR-01 contains a sequence similar to the 57 kbp insert in Ug99. **b** to **e,** Bar graphs show number of homozygous (orange) and heterozygous (blue) SNPs and MNPs called against the *Pgt*21-0 or Ug99 A, B and C haplotypes from Illumina read data for 43 *Pgt* isolates used for phylogenetic analysis. Read mapping patterns to each haplotype vary according to the presence or absence of either the A, B or C haplotypes in each isolate. Considering read mapping to the *Pgt*21-0 reference first; for *Pgt*21-0 and the other clonal Australian and South African isolates containing both A and B haplotypes, reads from the A nucleus will map to the A genome and reads from the B nucleus will map to the B genome. A very low number of homozygous SNPs are therefore detected that represent accumulated mutations as this lineage has evolved. For Ug99 and other A+C haplotype isolates in this clonal group, reads from the A nucleus will map to the A genome and again any new mutations appear as a small number of homozygous SNPs in the A genome data set. However, reads from the C genome can map to either the A or B genomes according to sequence similarity. Those reads that map to the B genome will give rise to homozygous SNPs representing divergence between the B and C genomes. However, reads that map to the A genome will give rise to heterozygous SNPs because A genome reads are already mapped to these regions. Thus, for A+C haplotype isolates, we see a small number of homozygous SNPs but a large number of heterozygous SNPs called against the A genome of *Pgt*21-0 (**b**), while the reverse is true for SNPs called against the B genome (**c**). Other isolates that are not hybrids derived from race 21 will have two nuclei that are neither A nor B, and reads from these can map to either the A or B genomes giving rise to high numbers of both heterozygous and homozygous SNPs on each haplotype. Thus, the variation in these patterns of heterozygous and homozygous SNPs on the different haplotypes are indicative of different hybrid relationships. The patterns for CZ-02,03 and PK-01,02 are consistent with them containing haplotype B, with many heterozygous but very few homozygous SNPs called on this haplotype, while IR-01 shows a similar pattern on haplotype C, again indicating that it contains a very similar haplotype.

List of other supplementary tables supplied as additional files in the publication (https://www.biorxiv.org/content/10.1101/692640v1)

Supplementary Table 1. Virulence reactions and pathotype assignments of *Pgt* isolates in the Ug99 lineage. Scores are reported based on the North American wheat differential set (excel file).

Supplementary Table 5. Gene synteny output (excel file)

Supplementary Table 7. Summary of karyon assignment before breaking chimeric contigs in Ug99 and *Pgt*21-0 (excel file)

Supplementary Table 8. List of chimeric contigs and breakpoints (excel file)

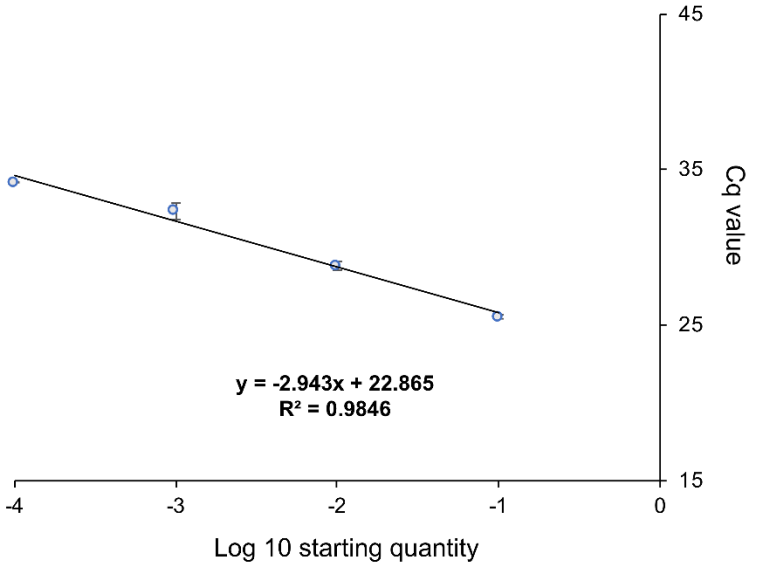
Supplementary Table 10. Physical linkage of phase swap contigs in the *Pgt*21-0 assembly to contigs of the same or alternate haplotype within bin or chromosome calculated from Hi-C data (excel file)

Supplementary Table 12. Contigs assigned to chromosomes (excel file)

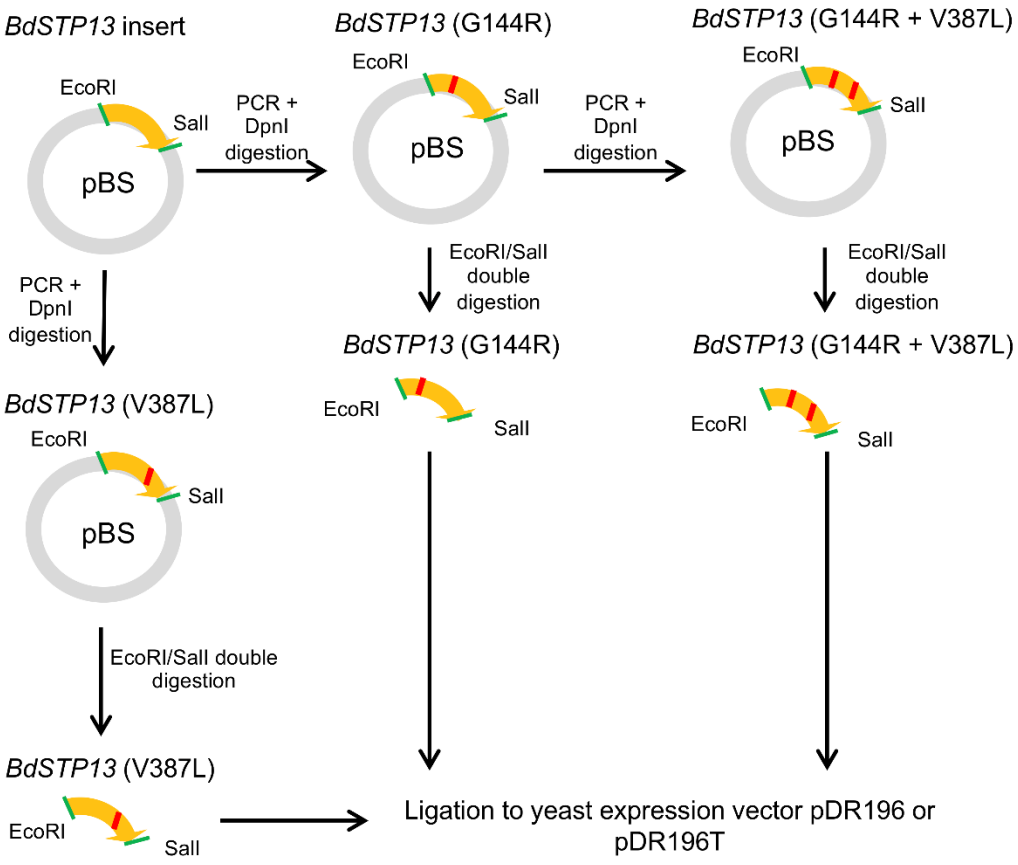
Supplementary Table 16. Metadata for RNAseq libraries of *Pgt*21-0 used for training gene models in the annotation pipeline (excel file).

Supplementary Table 17. Metadata genome coverages after mapping Illumina reads to *Pgt*21-0 and Ug99 references (excel file)

Appendix 2: Supplementary Materials of Chapter 3



Supplemental Figure 3.1. Quantification PCR efficiency test of primers targeting the amplification of *BdSTP13* transcript.
Reverse-transcription quantitative PCR (RT-qPCR) assay performed using cDNA template of WT Bd21-3 and primers MF125, MF126. Y-axis reflects the quantification cycle (Cq). Each data point represents the mean ± SD of three technical replicates.



Supplemental Figure 3.2. Workflow and schematic diagram for generating plasmid constructs of *BdSTP13* and its mutants by site-directed mutagenesis.
The nucleotide acid changes corresponding to G144R and V387L in single mutant (*BdSTP13*^{G144R}, *BdSTP13*^{V387L}) or double mutant (*BdSTP13*$^{G144R+V387L}$) were introduced into *BdSTP13* by PCR-mediated site-directed mutagenesis. The backbone plasmid pBluescript SK (-) (pBS) is shown as a grey circle, and *BdSTP13* cDNA insert is shown by a yellow arrow. The red short lines within the yellow arrow represent mutation positions. The green short lines (ends of *BdSTP13* insert) represent restriction enzyme sites (EcoRI and SalI). The fragment of single mutant (*BdSTP13*^{G144R}, *BdSTP13*^{V387L}) or double mutant (*BdSTP13*$^{G144R+V387L}$) were subsequently ligated to a yeast expression vector pDR196 or pDR196T.

Supplemental Table 3.1 and 3.2 in this chapter were excel worksheets provided in the following link:

Chapter 3 Supplemental Table 3.1 T-DNA insertional mutants used in this study.xlsx

Chapter 3 Supplemental Table 3.2. RNAseq for gens in the prioritized list.xlsx

Appendix 3: Loss of *AvrSr50* by somatic exchange in stem rust leads to virulence for *Sr50* resistance in wheat[2]

This article originally published in *Science* under the Science Journals Default License. Paper could be accessed at URL: https://www.ncbi.nlm.nih.gov/pubmed/29269475. Authorization letter to reuse this material in the dissertation was also included here.

Abstract

Race-specific resistance genes protect the global wheat crop from stem rust disease caused by *Puccinia graminis* f. sp. *tritici* (*Pgt*) but are often overcome owing to evolution of new virulent races of the pathogen. To understand virulence evolution in *Pgt*, we identified the protein ligand (AvrSr50) recognized by the Sr50 resistance protein. A spontaneous mutant of *Pgt* virulent to *Sr50* contained a 2.5 mega–base pair loss-of-heterozygosity event. A haustorial secreted protein from this region triggers *Sr50*-dependent defense responses in planta and interacts directly with the Sr50 protein. Virulence alleles of *AvrSr50* have arisen through DNA insertion and sequence divergence, and our data provide molecular evidence that in addition to sexual recombination, somatic exchange can play a role in the emergence of new virulence traits in *Pgt*.

Contribution to this work:

In this publication I was responsible to clone and characterize a fungal effector gene *AvrSr50* and its allelic variation (including structure variation) in several *Pgt* isolates including North American races MCCFC, SCCL and QCMJC and an African isolate TTKSK (Ug99), utilizing plasmid transformation and molecular cloning followed by Sanger sequencing (Figure 2E, Supplemental Figure 9). Using this information, I assessed

[2] Chen, J.†, Upadhyaya, N.M.†, Ortiz, D.†, Sperschneider, J., **Li, F.,** Bouton, C., Breen, S., Dong, C., Xu, B., Zhang, X., Mago, R., Newell, K., Xia, X., Bernoux, M., Taylor, J.M., Steffenson, B., Jin, Y., Zhang, P., Kanyuka, K., Figueroa, M., Ellis, J.G., Park, R.F., Dodds, P.N., 2017. Loss of *AvrSr50* by somatic exchange in stem rust leads to virulence for *Sr50* resistance in wheat. Science 358, 1607–1610. https://doi.org/10.1126/science.aao4810
†These authors contributed equally to the manuscript.

the variation in the *AvrSr50* locus in a set of global stem rust isolates and determined the likelihood of the pathogen to overcome resistance provided by the gene *Sr50*.

Authorization letter:

Appendix 4: Detection of Race-Specific Resistance Against *Puccinia coronata* f. sp. *avenae* in *Brachypodium* Species[3]

Abstract

Oat crown rust caused by *Puccinia coronata* f. sp. *avenae* is the most destructive foliar disease of cultivated oat. Characterization of genetic factors controlling resistance responses to *Puccinia coronata* f. sp. *avenae* in nonhost species could provide new resources for developing disease protection strategies in oat. We examined symptom development and fungal colonization levels of a collection of *Brachypodium distachyon* and *B. hybridum* accessions infected with three North American *P. coronata* f. sp. *avenae* isolates. Our results demonstrated that colonization phenotypes are dependent on both host and pathogen genotypes, indicating a role for race-specific responses in these interactions. These responses were independent of the accumulation of reactive oxygen species. Expression analysis of several defense-related genes suggested that salicylic acid and ethylene-mediated signaling but not jasmonic acid are components of resistance reaction to *P. coronata* f. sp. *avenae*. Our findings provide the basis to conduct a genetic

[3] Reprinted from Omidvar, V., Dugyala, S., **Li, F.,** Rottschaefer, S.M., Miller, M.E., Ayliffe, M., Moscou, M.J., Kianian, S.F., Figueroa, M., 2018. Detection of Race-Specific Resistance Against *Puccinia coronata* f. sp. *avenae* in Brachypodium Species. Phytopathology 108, 1443–1454. https://doi.org/10.1094/PHYTO-03-18-0084-R

inheritance study to examine whether effector-triggered immunity contributes to nonhost resistance to *P. coronata* f. sp. *avenae* in *Brachypodium* spp.

Contribution to this work:

In this publication I assisted with inoculation experiments and participated in characterizing expression profiling of a sugar transporter gene *BdSTP13* in *Brachypodium distachyon*. I also contributed to the overall editing of the manuscript.

Appendix 5: Pushing the boundaries of resistance: insights from *Brachypodium*-rust interactions[4]

Abstract

The implications of global population growth urge transformation of current food and bioenergy production systems to sustainability. Members of the family Poaceae are of particular importance both in food security and for their applications as biofuel substrates. For centuries, rust fungi have threatened the production of valuable crops such as wheat, barley, oat, and other small grains; similarly, biofuel crops can also be susceptible to these pathogens. Emerging rust pathogenic races with increased virulence and recurrent rust epidemics around the world point out the vulnerability of monocultures. Basic research in plant immunity, especially in model plants, can make contributions to understanding plant resistance mechanisms and improve disease management strategies. The development of the grass *Brachypodium distachyon* as a genetically tractable model for monocots, especially temperate cereals and grasses, offers the possibility to overcome the experimental challenges presented by the genetic and genomic complexities of economically valuable crop plants. The numerous resources and tools available in *Brachypodium* have opened new doors to investigate the underlying molecular and genetic bases of plant–microbe interactions in grasses and evidence demonstrating the applicability and advantages of working with *B. distachyon* is increasing. Importantly, several interactions between *B. distachyon* and devastating plant pathogens, such rust fungi, have been examined in the context of non-host resistance. Here, we discuss the use of *B.*

[4] Figueroa, M., Castell-Miller, C.V., **Li, F.,** Hulbert, S.H., Bradeen, J.M., 2015. Pushing the boundaries of resistance: insights from Brachypodium-rust interactions. Frontiers in plant science 6: 558. https://doi.org/10.3389/fpls.2015.00558

distachyon in these various pathosystems. Exploiting *B. distachyon* to understand the mechanisms underpinning disease resistance to non-adapted rust fungi may provide effective and durable approaches to fend off these pathogens. The close phylogenetic relationship among *Brachypodium* spp. and grasses with industrial and agronomic value support harnessing this model plant to improve cropping systems and encourage its use in translational research.

Contribution to this work:

In this publication I participated in writing the section "*Brachypodium distachyon*: A Model System that Defies Rust Diseases". I also contributed to the overall editing of the manuscript.

Appendix 6: *Puccinia coronata* f. sp. *avenae*: a threat to global oat production[5]

This article originally published in *Molecular Plant Pathology* at URL: https://bsppjournals.onlinelibrary.wiley.com/doi/abs/10.1111/mpp.12608. Authorization letter to reuse this material in the dissertation was also included here.

Summary

Puccinia coronata f. sp. *avenae* (*Pca*) causes crown rust disease in cultivated and wild oat (*Avena* spp.). The significant yield losses inflicted by this pathogen make crown rust the most devastating disease in the oat industry. *Pca* is a basidiomycete fungus with an obligate biotrophic lifestyle, and is classified as a typical macrocyclic and heteroecious fungus. The asexual phase in the life cycle of *Pca* occurs in oat, whereas the sexual phase takes place primarily in *Rhamnus* species as the alternative host. Epidemics of crown rust happens in areas with warm temperatures (20–25 °C) and high humidity. Infection by the pathogen leads to plant lodging and shrivelled grain of poor quality.

Disease symptoms: Infection of susceptible oat varieties gives rise to orange–yellow round to oblong uredinia (pustules) containing newly formed urediniospores. Pustules vary in size and can be larger than 5 mm in length. Infection occurs primarily on the surfaces of leaves, although occasional symptoms develop in the oat leaf sheaths and/or floral structures, such as awns. Symptoms in resistant oat varieties vary from flecks to small pustules, typically accompanied by chlorotic halos and/or necrosis. The pycnial and aecial stages are mostly present in the leaves of *Rhamnus* species, but occasionally symptoms can also be observed in petioles, young stems and floral structures. Aecial structures display a characteristic hypertrophy and can differ in size, occasionally reaching more than 5 mm in diameter.

[5] Nazareno, E.S., **Li, F.,** Smith, M., Park, R.F., Kianian, S.F., Figueroa, M., 2018. *Puccinia coronata* f. sp. *avenae*: a threat to global oat production. Mol. Plant Pathol. 19, 1047–1060. https://doi.org/10.1111/mpp.12608

Taxonomy: *Pca* belongs to the kingdom Fungi, phylum Basidiomycota, class Pucciniomycetes, order Pucciniales and family Pucciniaceae.

Host range: *Puccinia coronata sensu lato* can infect 290 species of grass hosts. *Pca* is prevalent in all oat-growing regions and, compared with other cereal rusts, displays a broad telial host range. The most common grass hosts of *Pca* include cultivated hexaploid oat (*Avena sativa*) and wild relatives, such as bluejoint grass, perennial ryegrass and fescue. Alternative hosts include several species of *Rhamnus*, with *R. cathartica* (common buckthorn) as the most important alternative host in Europe and North America.

Control: Most crown rust management strategies involve the use of rust-resistant crop varieties and the application of fungicides. The attainment of the durability of resistance against *Pca* is difficult as it is a highly variable pathogen with a great propensity to overcome the genetic resistance of varieties. Thus, adult plant resistance is often exploited in oat breeding programmes to develop new crown rust-resistant varieties.

Contribution to this work:

In this publication I participated in writing the section "*Brachypodium distachyon*: A Model System that Defies Rust Diseases". I was responsible to write sections on the economic losses due to oat crown rust and non-host resistance to rust pathogens using Brachypodium as a genetic source. I also contributed to the overall editing of the manuscript.

Authorization letter:

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Aug 26, 2019

This Agreement between Feng Li ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4656651200232
License date	Aug 26, 2019
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Molecular Plant Pathology
Licensed Content Title	Puccinia coronata f. sp. avenae: a threat to global oat production
Licensed Content Author	Melania Figueroa, Shahryar F. Kianian, Robert F. Park, et al
Licensed Content Date	Dec 10, 2017
Licensed Content Volume	19
Licensed Content Issue	5
Licensed Content Pages	14
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Abstract
Will you be translating?	No
Order reference number	oat crown rust paper review
Title of your thesis / dissertation	Understanding wheat stem rust virulence and non-host resistance against stem rust
Expected completion date	Oct 2019
Expected size (number of pages)	150
Requestor Location	Feng Li 1204 Fifield Street SAINT PAUL, MN 55108 United States Attn: Feng Li
Publisher Tax ID	EU826007151
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright